

AFIT/DS/ENG/97-05

PARAMETER ESTIMATION FOR  
REAL FILTERED SINUSOIDS

DISSERTATION  
Daniel R. Zahirniak  
Major, USAF

AFIT/DS/ENG/97-05

19971022 075

Approved for public release; distribution unlimited

DTIC QUALITY INSPECTED 8

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the Department of Defense or the U. S. Government.

PARAMETER ESTIMATION FOR  
REAL FILTERED SINUSOIDS

DISSERTATION

Presented to the Faculty of the Graduate School of Engineering  
of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the  
Requirements for the Degree of  
Doctor of Philosophy

Daniel R. Zahirniak, B.S.E.E, M.S.E.E  
Major, USAF

September, 1997

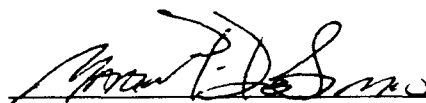

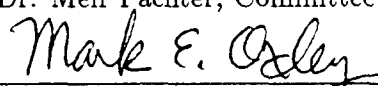
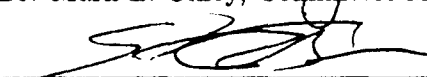
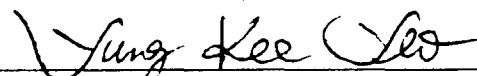
Approved for public release; distribution unlimited

PARAMETER ESTIMATION FOR  
REAL, FILTERED SINUSOIDS

Daniel R. Zahirniak, B.S.E.E., M.S.E.E.

Major, USAF

Approved:

	<u>19 Sept 97</u>
Dr. Martin P. DeSimio, Committee Chairman	
	<u>18 Sept 97</u>
Dr. Meir Pachter, Committee Member	
	<u>18 Sept 97</u>
Dr. Mark E. Oxley, Committee Member	
	<u>18 Sept 97</u>
Dr. Steven K. Rogers, Committee Member	
	<u>19 Sept 97</u>
Dr. Yung Kee Yeo, Dean's Representative	



Dr. Robert A. Calico, Jr., Dean

## Table of Contents

	Page
List of Abbreviations . . . . .	ix
List of Symbols . . . . .	xi
List of Figures . . . . .	xiv
List of Tables . . . . .	xvi
Acknowledgements . . . . .	xvii
Abstract . . . . .	xviii
I. Introduction . . . . .	1-1
1.1 Problem . . . . .	1-1
1.2 Application . . . . .	1-2
1.3 Scope . . . . .	1-3
1.4 Contributions . . . . .	1-4
1.5 Overview . . . . .	1-6
II. Literature Review: Sinusoidal Frequency Estimation . . . . .	2-1
2.1 Introduction . . . . .	2-1
2.2 Single Sinusoid Frequency Estimation . . . . .	2-1
2.3 Multiple Sinusoid Frequency Estimation . . . . .	2-5
2.3.1 Algebraic Techniques . . . . .	2-7
2.3.2 Iterative Filtering . . . . .	2-11
2.3.3 Maximum Likelihood . . . . .	2-14
2.4 Summary . . . . .	2-16

	Page
III. Theory: Sinusoidal Parameter Estimation . . . . .	3-1
3.1 Introduction . . . . .	3-1
3.2 Filtered Data Model . . . . .	3-1
3.2.1 Input Signal . . . . .	3-1
3.2.2 Output Signal . . . . .	3-2
3.2.3 Model Development . . . . .	3-3
3.2.4 Section Summary . . . . .	3-5
3.3 Sinusoidal Parameter Estimation . . . . .	3-5
3.3.1 Estimation Background . . . . .	3-5
3.3.2 Estimation Accuracy . . . . .	3-7
3.3.3 Estimator Evaluation . . . . .	3-8
3.3.4 Maximum Likelihood (ML) Estimation . . . . .	3-8
3.3.5 Parameter Estimators . . . . .	3-10
3.3.6 ML Frequency Estimation . . . . .	3-11
3.3.7 Section Summary . . . . .	3-12
3.4 Linear Prediction (LP) Modeling . . . . .	3-17
3.4.1 Theoretical Background . . . . .	3-17
3.4.2 Construction of the LP General Linear Model . . . . .	3-18
3.4.3 Noise Effects in the Measurements . . . . .	3-18
3.4.4 Noise Effects in the Observation Matrix . . . . .	3-20
3.4.5 Imposition of Sinusoidal Constraints . . . . .	3-21
3.4.6 Bounding the Estimation Error . . . . .	3-22
3.4.7 Section Summary . . . . .	3-23
3.5 LP Coefficient Estimation: Fixed Point Concept . . . . .	3-25
3.5.1 Theoretical Background . . . . .	3-25
3.5.2 Metric Space Definition . . . . .	3-26
3.5.3 Iterative Least Squares Fixed Point Mapping . . . . .	3-26

	Page
3.5.4 Iterative Total Least Squares Fixed Point Mapping .	3-32
3.5.5 Section Summary . . . . .	3-39
3.6 LP Modeling/ ML Frequency Estimation Relationship . . . .	3-39
3.6.1 ML Objective Function Redefined . . . . .	3-43
3.6.2 Determination of the Basis Vectors . . . . .	3-45
3.6.3 Relationship to LP Model . . . . .	3-46
3.6.4 Section Summary . . . . .	3-47
3.7 LP Coefficient Estimation: Objective Function Minimization .	3-47
3.7.1 Iterative Exact Gradient Descent (IEGD) . . . . .	3-47
3.7.2 Iterative Generalized Least Squares (IGLS) . . . . .	3-49
3.7.3 Iterative Quadratic Maximum Likelihood (IQML) .	3-51
3.7.4 Algorithm Comparison . . . . .	3-52
3.7.5 Section Summary . . . . .	3-55
3.8 Bounding the Estimation Error . . . . .	3-55
3.8.1 Theory . . . . .	3-56
3.8.2 LP Coefficients . . . . .	3-57
3.8.3 Sinusoidal Frequencies . . . . .	3-60
3.8.4 Section Summary . . . . .	3-63
3.9 Summary . . . . .	3-65
IV. Application: Digital Electronic Warfare Receiver . . . . .	4-1
4.1 Introduction . . . . .	4-1
4.2 Digital Receiver Architecture . . . . .	4-1
4.2.1 Real Data Model . . . . .	4-4
4.2.2 Complex Data Model . . . . .	4-6
4.2.3 Receiver Implementation . . . . .	4-8
4.3 Parameter Estimation . . . . .	4-12
4.3.1 Estimation Accuracy . . . . .	4-12

	Page
4.3.2 ML Estimation . . . . .	4-14
4.3.3 LP Modeling . . . . .	4-14
4.3.4 LP Model-ML Frequency Estimation Relationship .	4-18
4.3.5 LP Coefficient Estimation . . . . .	4-19
4.3.6 Complex Model Approximation . . . . .	4-21
4.3.7 Bounding the Estimation Error . . . . .	4-25
4.3.8 Simplified Single Sinusoidal Estimator . . . . .	4-26
4.4 Summary . . . . .	4-30
V. Conclusion . . . . .	5-1
5.1 Introduction . . . . .	5-1
5.2 Theoretical Contributions . . . . .	5-1
5.3 Applied Contributions . . . . .	5-4
5.4 Summary . . . . .	5-6
Appendix A. Vector-Matrix Differentiation . . . . .	A-1
A.1 Definitions . . . . .	A-1
A.2 Matrix-Vector Product Derivative . . . . .	A-1
A.3 Matrix Product Derivative . . . . .	A-2
A.4 Inverse Derivative . . . . .	A-2
A.5 Vector-Matrix-Vector Product Derivative . . . . .	A-3
Appendix B. Maximizing ML Objective Functions . . . . .	B-1
B.1 Real Data ML Maximization . . . . .	B-1
B.2 Complex Data ML Maximization . . . . .	B-2
Appendix C. Linear Prediction Coefficient Constraints . . . . .	C-1
C.1 Real Sinusoids . . . . .	C-1
C.1.1 FBLP Constraints . . . . .	C-1



	Page
C.1.2 Symmetry Constraints . . . . .	C-2
C.2 Complex Sinusoids . . . . .	C-3
C.2.1 Forward Backward Linear Prediction . . . . .	C-3
C.2.2 Symmetry Constraints . . . . .	C-4
Appendix D. Fixed Point Theory . . . . .	D-1
D.1 Definitions . . . . .	D-1
D.1.1 Linear Space . . . . .	D-1
D.1.2 Euclidean N-Space . . . . .	D-1
D.1.3 Metric Space . . . . .	D-2
D.1.4 Complete Metric Space . . . . .	D-2
D.1.5 Banach Space . . . . .	D-2
D.1.6 Open and Closed Sets . . . . .	D-2
D.1.7 Bounded and Compact Sets . . . . .	D-2
D.1.8 Continuous and Bounded Functions . . . . .	D-3
D.1.9 Fixed Point . . . . .	D-3
D.2 Fixed Point Existence Theorems . . . . .	D-3
D.2.1 Schauder's Fixed Point Theorem . . . . .	D-3
D.2.2 Brouwer's Fixed Point Theorem . . . . .	D-3
Appendix E. Objective Function Gradients . . . . .	E-1
E.1 ML Frequency Objective Function Gradient . . . . .	E-1
E.2 ML Objective Function Gradient . . . . .	E-2
Appendix F. Short Time Fourier Transform/Filter Bank Equivalence . .	F-1
Appendix G. Complex Random Variables and Vectors . . . . .	G-1
G.1 Random Variables . . . . .	G-1
G.2 Random Vectors . . . . .	G-2
G.3 Wide Sense Stationary . . . . .	G-4

	Page
Appendix H.      Complex Normally Distributed Random Vectors . . . . .	H-1
Appendix I.      Alternate Architecture . . . . .	I-1
I.1    New Architecture . . . . .	I-1
I.2    Previous Architecture . . . . .	I-2
I.3    Architecture Relationship . . . . .	I-4
Appendix J.      Phase Noise Analysis . . . . .	J-1
Bibliography . . . . .	BIB-1
Vita . . . . .	VITA-1

# *List of Abbreviations*

Abbreviation	Page
(EW) - Electronic Warfare . . . . .	1-2
(IF) - Intermediate Frequency . . . . .	1-2
(SNR) - Signal-to-Noise Ratio . . . . .	1-2
(DFT) - Discrete Fourier Transform . . . . .	2-1
(WPA) -Weighted Phase Averaging . . . . .	2-2
(IPA) - Iterative Phase Averaging . . . . .	2-3
(FFT) - Fast Fourier Transform . . . . .	2-5
(LP) - Linear Prediction . . . . .	2-7
(FBLP) - Forward-Backward Linear Prediction . . . . .	2-7
(LS) - Least Squares . . . . .	2-8
(PE) - Principal Eigenvalue . . . . .	2-8
(TLS) - Total Least Squares . . . . .	2-10
(SVD) - Singular Valued Decomposition . . . . .	2-10
(IFA) - Iterative Filtering Algorithm . . . . .	2-12
(PF) - Parametric Filtering . . . . .	2-12
(IQML) - Iterative Quadratic Maximum Likelihood . . . . .	2-15
(PDF) - Probability Density Function . . . . .	3-6
(CRLB) - Cramer Rao Lower Bound . . . . .	3-7
(MSE) - Mean Square Error . . . . .	3-8
(ML) - Maximum Likelihood . . . . .	3-8
(LP) - Linear Prediction . . . . .	3-17
(WLS) - Weighted Least Squares . . . . .	3-19
(ILS) - Iterative Least Squares . . . . .	3-25
(ITLS) - Iterative Total Least Squares . . . . .	3-25
(ILS) - Iterative Least Squares . . . . .	3-30

Abbreviation	Page
(ITLS) - Iterative Total Least Squares . . . . .	3-37
(IGLS) - Iterative Generalized Least Squares . . . . .	3-47
(IQML) - Iterative Total Least Squares . . . . .	3-47
(IEGD) - Iterative Exact Gradient Descent . . . . .	3-48
(AFRL/WL) - Air Force Research Laboratory/Wright Laboratory . . . . .	4-1
(UDFT) - Uniform Discrete Fourier Transform . . . . .	4-1
(ADC) - Analog to Digital Converter . . . . .	4-1
(BW) - Bandwidth . . . . .	4-2
(STFT) - Short Time Fourier Transform . . . . .	4-3
(IDFT) - Inverse Discrete Fourier Transform . . . . .	4-8
(WSS) - Wide Sense Stationary . . . . .	G-4

# *List of Symbols*

Symbol	Page
$P$ - Number of filtered sinusoids . . . . .	2-1
$s[m]$ - Linear sum of $P$ filtered sinusoids . . . . .	2-1
$w[m]$ - Filtered noise . . . . .	2-1
$y[m]$ - Linear sum of $P$ filtered sinusoids in noise . . . . .	2-1
$\{b_k\}$ - Filtered sinusoid amplitudes . . . . .	2-1
$\{\phi_k\}$ - Filtered sinusoid phases . . . . .	2-1
$\{f_k\}$ - Filtered sinusoid frequencies . . . . .	2-1
$\omega$ - Radian Frequency . . . . .	2-2
$a[p]$ - Linear prediction coefficients . . . . .	2-7
$A(z)$ - Linear prediction polynomial . . . . .	2-7
$\mathbf{a}$ - Vector of linear prediction coefficients . . . . .	2-7
$\mathbf{a}_o$ - Vector of truncated linear prediction coefficients . . . . .	2-7
$Y_F$ - Forward observation matrix . . . . .	2-7
$Y_B$ - Backward observation matrix . . . . .	2-7
$Y_{FB}$ - Forward, backward observation matrix . . . . .	2-8
$S_{FB}$ - Forward, backward signal matrix . . . . .	2-8
$e[m]$ - Model error . . . . .	2-12
$B$ - Matrix of sinusoidal constraints . . . . .	2-14
$\alpha$ - Vector of constrained LP coefficients . . . . .	2-14
$Y_C$ - Constrained observation matrix . . . . .	2-14
$J(\mathbf{a})$ - LP coefficient objective function . . . . .	2-14
$A$ - Matrix of LP coefficients . . . . .	2-14
$x[m]$ - Linear sum of $P$ sinusoids in independent noise . . . . .	3-1
$\nu[m]$ - Linear sum of $P$ sinusoids prior to filtering . . . . .	3-1
$\eta[m]$ - Zero-mean, independent, normally distributed noise . . . . .	3-1

Symbol	Page
$N$ - Length of filter . . . . .	3-2
$\{h[n]\}$ - Set of filter coefficients . . . . .	3-2
$H(e^{j\omega})$ - Frequency transfer function . . . . .	3-2
$\mathbf{y}$ - Vector of sinusoids in noise . . . . .	3-3
$\mathbf{s}$ - Vector of sinusoids . . . . .	3-3
$\mathbf{w}$ - Vector of noise . . . . .	3-3
$K_{\mathbf{w}}$ - Filtered noise covariance matrix . . . . .	3-3
$\Lambda$ - Deterministic signal matrix . . . . .	3-4
$\mathbf{b}$ - Vector of scaling coefficients . . . . .	3-5
$\boldsymbol{\theta}$ - Vector of parameters . . . . .	3-5
$E\{\star\}$ - Expected value operator . . . . .	3-6
$K_{\hat{\boldsymbol{\theta}}}$ - Estimate covariance matrix . . . . .	3-6
$p(\mathbf{y}; \boldsymbol{\theta})$ - Parameterized probability density function . . . . .	3-6
$\mathbf{f}$ - Vector of frequencies . . . . .	3-6
$\mathcal{F}(\boldsymbol{\theta})$ - Fisher information matrix . . . . .	3-7
$\mathbf{m}_{\mathbf{y}}$ - Mean vector of observations . . . . .	3-7
$J(\boldsymbol{\theta})$ - ML parameter objective function . . . . .	3-10
$J(\mathbf{f})$ - ML frequency objective function . . . . .	3-10
$G(\mathbf{a})$ - Cholesky decomposition of LP model covariance matrix . . . . .	3-21
$\mathcal{S}_L(0, k)$ - Set of $L$ -dimensional vectors of finite length . . . . .	3-26
$d(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j)$ - Distance function for two vectors . . . . .	3-26
$\mathcal{L}_{LS}$ - Least squares mapping function . . . . .	3-27
$\mathcal{L}_{ILS}$ - Iterative least squares mapping function . . . . .	3-30
$\mathcal{L}_{TLS}$ - Total least squares mapping function . . . . .	3-36
$\mathcal{L}_{ITLS}$ - Iterative total least squares mapping function . . . . .	3-37
$J(\boldsymbol{\alpha})$ - Constrained form of LP objective function . . . . .	3-47
$\hat{K}_{\hat{\boldsymbol{\alpha}}}$ - Covariance matrix for LP coefficients . . . . .	3-60

Symbol	Page
$V\{*\}$ - Variance operator . . . . .	3-60
$\tilde{h}_k[n]$ - Complex filter coefficients for $k^{th}$ EW bandpass filter . . . . .	4-3
$R$ - Decimation Rate . . . . .	4-4
$\mathbf{h}$ - Vector of filter coefficients . . . . .	4-5
$H$ - Filter matrix . . . . .	4-5
$\Lambda_{\mathcal{R}}$ - Deterministic decimated real signal matrix . . . . .	4-5
$\tilde{y}[m]$ - Sum of complex sinusoids in noise . . . . .	4-6
$\tilde{w}[m]$ - Complex filtered noise . . . . .	4-7
$w_r[m]$ - Real part of complex noise . . . . .	4-7
$w_i[m]$ - Imaginary part of complex noise . . . . .	4-7
$\tilde{\mathbf{y}}$ - Vector of complex filtered sinusoids in noise . . . . .	4-7
$\mathbf{y}_r$ - Vector of real components of complex filtered sinusoids in noise . . . . .	4-7
$\mathbf{y}_i$ - Vector of imaginary components of complex filtered sinusoids in noise . . . . .	4-7
$\Lambda_{\mathcal{C}}$ - Deterministic signal matrix for the complex model . . . . .	4-8
$\tilde{s}[m]$ - Sum of complex sinusoids . . . . .	4-14
$\tilde{a}[p]$ - Complex LP coefficients . . . . .	4-15
$\mathbf{a}_r$ - Vector of real components of LP coefficients . . . . .	4-15
$\mathbf{a}_i$ - Vector of imaginary components of LP coefficients . . . . .	4-15
$A_r$ - Matrix of real components of LP coefficients . . . . .	4-16
$A_i$ - Matrix of imaginary components of LP coefficients . . . . .	4-16
$A_{\mathcal{C}}$ - Matrix of coefficients for the complex LP model . . . . .	4-16
$\tilde{A}$ - Matrix of complex LP coefficients . . . . .	4-29

## *List of Figures*

Figure	Page
1.1. Electronic Warfare Receiver Block Diagram . . . . .	1-2
1.2. Typical Radar Signal . . . . .	1-3
2.1. WPA Frequency Estimation Accuracy . . . . .	2-4
2.2. IPA Frequency Estimation Accuracy . . . . .	2-6
2.3. LS Frequency Estimation Accuracy . . . . .	2-9
2.4. PE Frequency Estimation Accuracy . . . . .	2-10
2.5. TLS Frequency Estimation Accuracy . . . . .	2-11
2.6. IFA Frequency Estimation Accuracy . . . . .	2-13
2.7. PF Frequency Estimation Accuracy . . . . .	2-15
2.8. IQML Frequency Estimation Accuracy . . . . .	2-16
3.1. CRLB for Sinusoids in Noise . . . . .	3-9
3.2. Maximizing ML Frequency Objective Function . . . . .	3-13
3.3. ML Frequency Estimation Accuracy . . . . .	3-14
3.4. ML Amplitude Estimation Accuracy . . . . .	3-15
3.5. ML Phase Estimation Accuracy . . . . .	3-16
3.6. CRLB for LP Coefficients . . . . .	3-24
3.7. ILS Fixed Point Convergence . . . . .	3-33
3.8. ILS Frequency Estimation Accuracy-One Sinusoid . . . . .	3-34
3.9. ILS Frequency Estimation Accuracy-Two Sinusoids . . . . .	3-35
3.10. ITLS Fixed Point Convergence . . . . .	3-40
3.11. ITLS Frequency Estimation Accuracy-One Sinusoid . . . . .	3-41
3.12. ITLS Frequency Estimation Accuracy-Two Sinusoids . . . . .	3-42
3.13. IEGD Algorithm Convergence . . . . .	3-50
3.14. LP Estimator Minimum Error Comparison . . . . .	3-53



Figure	Page
3.15. LP Estimator Frequency Estimation Comparison . . . . .	3-54
3.16. Variance Estimation of LP Coefficients . . . . .	3-61
3.17. Frequency Confidence Intervals . . . . .	3-64
4.1. EW Digital Receiver Architecture . . . . .	4-2
4.2. UDFT Filter Bank Polyphase Implementation . . . . .	4-9
4.3. Nonmaximally Decimated Polyphase UDFT Filter Bank . . . . .	4-10
4.4. CRLB for Real and Complex Data Models . . . . .	4-13
4.5. ILS Estimation Accuracy for Single Sinusoid . . . . .	4-22
4.6. LS Frequency Estimation Accuracy Across Bandwidth . . . . .	4-23
4.7. ILS Frequency Estimation Accuracy for Two Difficult Sinusoids . . . . .	4-24
4.8. ILS Frequency Estimation Accuracy for Two Typical Sinusoids . . . . .	4-25
4.9. Variance Estimation of Complex LP Coefficients . . . . .	4-27
4.10. Frequency Confidence Intervals . . . . .	4-28
I.1. Conventional Nonmaximally Decimated UDFT Filter Bank . . . . .	I-6

*List of Tables*

Table	Page
4.1. Filter Critical Frequencies . . . . .	4-3
B.1. Derivatives and Gradients . . . . .	B-4

### *Acknowledgements*

There are many people who deserve a lot thanks for helping my succeed in the program here at AFIT. First thanks must go to my advisor, Dr. DeSimio for allowing me to pursue my research in the field of Signal Processing. Not only did he save me from the "black-hole" of electromagnetics research, he was also a good source for knowledge in the field and my severest critic when reviewing the dissertation. Without his effort, things could have been a lot more difficult.

Thanks also goes to Dr. Pachter for the personal and technical guidance he provided during years. He helped me to avoid the endless pitfalls which can be encountered when conducting research. The seeds for most of the ideas and contributions resulting from this research were planted by him. I am very thankful for his instruction and direction.

I would also like to thank Dave Sharpin for the critical assistance and instruction he provided in the application phase of my research.

Finally, and most importantly, I would like to thank my wife, Annette, who endured these past three years along with me and without whom I could not have completed this program. Her endless support was a constant light at the end of the tunnel.

Daniel R. Zahirniak

*Abstract*

Estimating the parameters of filtered sinusoids in noise from a finite number of observations has wide Air Force, DoD, and commercial interest. Whether constructing models for time series analysis or estimating the velocity of moving targets, the ability to accurately estimate the frequency, amplitude, and phase parameters of sinusoids is of paramount importance. This research develops theoretical methods of analyzing filtered sinusoids in noise and demonstrates their effectiveness on the problem of pulsed sinusoid parameter estimation for Electronic Warfare (EW) applications. Specifically, within the context of stochastic modeling, a new linear model, parameterized by a set of Linear Prediction (LP) coefficients, is derived for estimating the frequencies of filtered sinusoids. This model is an improvement over previous modeling techniques since the effects of the filter and the coefficients upon the noise statistics are properly accounted for during model development. In addition, the LP coefficients which minimize the squared error between the system model and the observations, are shown to be maximum likelihood coefficient estimates. Two methods of estimating the coefficients, based on an iterative least squares (ILS) and iterative total least squares (ITLS) solution to an over determined system of equations, are derived and shown to be fixed point mappings of the coefficients within the domain of allowable solutions. Application of this new linear model and the ILS estimator to the EW problem shows the ILS algorithm outperforms the current estimation techniques by providing optimal frequency estimates for multiple, filtered, pulsed sinusoids at lower signal to noise ratios. In addition, a bound for the estimation error of the LP coefficients and the frequencies, based on one set of observations, is derived and used to gauge the quality of a point estimate and establish associated confidence intervals. The results of this research, whether taken individually or collectively, represent new contributions to the theory of signal processing and parameter estimation and support the many applications requiring accurate parameter estimation of both complex and real filtered sinusoids in noise.

# PARAMETER ESTIMATION FOR REAL FILTERED SINUSOIDS

## *I. Introduction*

### *1.1 Problem*

This dissertation investigates the problem of estimating the parameters of a linear sum of filtered sinusoids and noise. Accurate estimation of the frequency, amplitude, and phase parameters of a signal containing a sum of sinusoids in noise is a common task in many applications in the field of applied science, engineering and statistics (39:407). For example, in the field of forecasting, it is often desired to decompose a time series into its main components of trend, irregular, cyclical and seasonal. This decomposition allows a model of the time series to be constructed for time-series analysis and prediction (23:1-16). Since a linear sum of sinusoids in zero-mean, normally distributed, white noise can be used to represent arbitrary time-series data by these components (76), accurate model construction relies on accurate estimation of the corresponding sinusoidal parameters.

Additionally, various types of radar systems require accurate frequency measurements to determine the relative velocity of a target and to separate moving targets from stationary objects (79:243-357). It is well known the electromagnetic wave reflected by an object moving in relation to the radar will be compressed in the direction of motion (74:464-466). This wave compression results in an apparent change in frequency, known as the Doppler effect, between the transmitted signal and the reflected signal (77:68-148). Since the amount frequency change is directly related to the relative velocity of the object, the object's velocity can be estimated by determining the difference between the carrier frequency of the transmitted signal and the frequency of the received signal (64:727-728). Thus, accurate frequency measurements are necessary for obtaining high resolution velocity measurements and discriminating between moving objects and clutter (62).

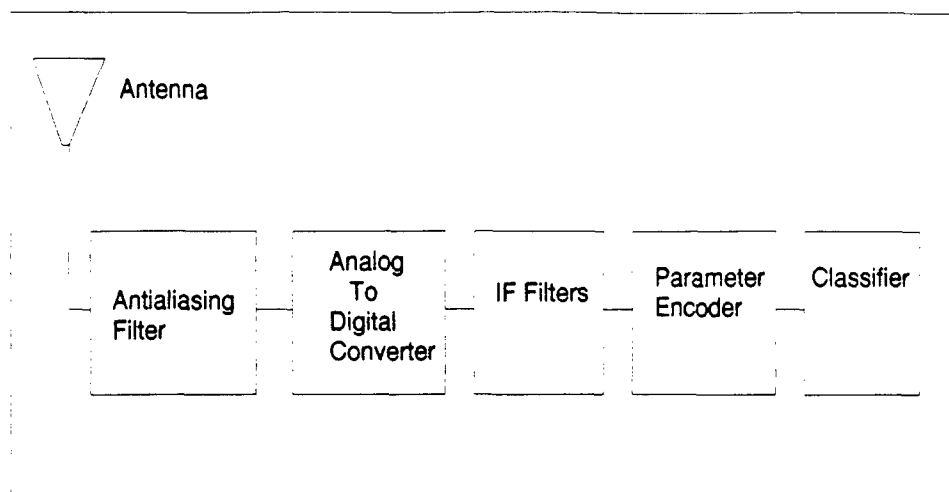


Figure 1.1 Electronic Warfare Receiver Block Diagram

## 1.2 Application

A critical military application which requires accurate sinusoidal parameter estimation involves Electronic Warfare (EW) receivers. The primary function of an EW receiver, shown by the block diagram in Figure 1.1, is to detect a pulsed radar signal, extract and encode signal information, and pass this data to a classification system for radar system identification (90:7-35).

The antenna and antialiasing filter are usually designed to intercept radar signals, modeled by pulsed sinusoids as depicted in Figure 1.2, typically in the frequency range 2 – 18 GHz (90:10). The analog-to-digital converter then digitizes the signal and passes the resulting discrete-time representation through a set of Intermediate Frequency (IF) filters. These IF filters limit the number of time-coincident sinusoids processed by the encoder and increase the Signal-to-Noise Ratio (SNR) of the sinusoids within the filter passband (90:13). The function of a typical encoder is to extract the sinusoidal parameter information present in the signal and provide point estimates of these parameters to the classifier. The classifier then identifies the radar system generating the signal by comparing these estimates with stored radar system signal descriptions.

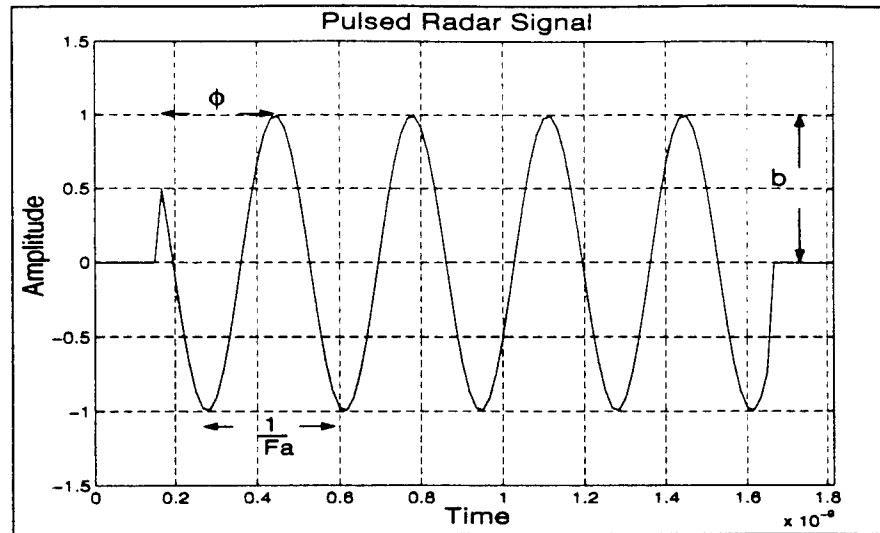


Figure 1.2 **Typical Radar Signal:** Digital representation of pulsed radar signal sampled at  $f_s = 60GHz$ . Signal parameters: Amplitude:  $b = 1$ , Frequency:  $F_a = 3GHz$ , Phase:  $\phi = \pi/3$ .

Thus, accurate estimates of the sinusoidal parameters are critical for both the construction of the classification system and for characterizing the waveform parameters of the pulsed sinusoids prior to radar system identification.

### 1.3 Scope

The primary focus of this dissertation is limited to investigating and developing algorithms which provide accurate point estimates of the parameters of real, filtered sinusoids in noise. During algorithm development, the amplitudes, frequencies and phases of the sinusoids are assumed to be deterministic and constant throughout the pulse. Though recent research has been conducted to estimate the parameters of chirp signals (41),(72), pulsed sinusoids still describe the majority of radar signals present and will provide a basis for examining more exotic waveforms in the future (90:10). Additionally, the characteristics of the EW receiver will assumed to be known and capable of being modeled as a finite impulse response, linear, time-invariant filter. It is the sinusoidal parameters which are unknown for this application. Furthermore, prior to filtering, the noise will be assumed to be a zero-mean, normally distributed, uncorrelated, wide sense stationary random pro-

cess with a known variance. Finally, the number of sinusoids present within any block of data will be assumed known prior to providing parameter estimates. Identification of the number of sinusoids present in a block of data is an area of research in its own right and can be accomplished prior to the employment of parameter estimation techniques (1), (4), (11), (18), (25), (41), (93).

#### *1.4 Contributions*

The results of this research, as delineated in Chapter III and Chapter IV, are highlighted via the original contributions below:

- Development of the true mathematical model for any digital system required to estimate the amplitudes, phases and frequencies of filtered sinusoids in noise. Most, if not all, digital signal processing systems implement a series of filters to condition the noisy data prior to estimating the sinusoidal parameters. The model developed as part of this research accounts for the effects of the filter in developing parameter estimators.
- Derivation of Maximum Likelihood (ML) estimators for the amplitude, phase and frequency parameters of filtered sinusoids in noise. By correctly accounting for the correlation in the noise due to the effects of the filter, an accurate model of the probability density function of the filtered signal is derived. ML estimation techniques are then used to provide parameter estimators. Simulations indicate that failure to incorporate the effects of the filter into the data model will lead to suboptimal parameter estimates.
- Construction of the true general linear model, parameterized by a set of Linear Prediction (LP) coefficients, for estimating the frequencies of filtered sinusoids in noise. By properly incorporating the effects of both the filter and the LP coefficients upon the noise, this new model represents the true linear model for estimating the LP coefficients and the sinusoidal frequencies from the measurements.
- Application of fixed point theory to the estimation of the coefficients for the true LP general linear model and development of the Iterative Least Squares (ILS) and



Iterative Total Least Squares (ITLS) fixed point mapping functions. Simulations indicate both methods provide minimum variance, unbiased estimates of the LP coefficients, and consequently, the sinusoidal frequencies, over a wide range of SNRs (98).

- Derivation of the exact relationship between the ML frequency objective function and the true LP general linear model for filtered sinusoids. Specifically, the set of LP coefficients which minimize the squared error defined by the LP general linear model are shown to provide ML frequency estimates and vice versa.
- Development of an exact ML estimator, termed the Iterative Exact Gradient Descent (IEGD) algorithm, for LP coefficient estimation. Simulations indicate the LP coefficients provided by the IEGD algorithm minimize the LP objective function thereby providing ML estimates of the LP coefficients and, consequently, the sinusoidal frequencies.
- Proof that the two most commonly used methods of estimating the LP coefficients, termed Iterative Generalized Least Squares (IGLS) and Iterative Quadratic Maximum Likelihood (IQML), are equivalent to the ILS and ITLS fixed point functional mappings, not minimization algorithms as widely accepted. Although these two estimators do not minimize the LP objective function, simulations indicate they suffer only a slight decrease in estimation accuracy, particularly at high SNRs.
- Derivation of a novel method for bounding the estimation error of point estimates of the LP coefficients and consequently, the frequencies, based strictly on one realization of the measurement vector. Simulations indicate the estimates of the measurement error can be used to establish confidence intervals for point estimates of both the LP coefficients and the frequencies.
- Construction of the data model which accurately describes the passage of real sinusoids through complex filters. The model shows the complex output of a band-limited filter can be decimated at twice the rate as the corresponding real output without a loss in estimation accuracy.

- Derivation of a complex form of the LP general linear model for estimating the frequencies of complex, filtered sinusoids in noise. By properly incorporating the effects of both the filter and the LP coefficients upon the noise, the complex, general linear model developed represents the true linear model for estimating the complex LP coefficients (97).
- Development of a multirate, channelized, EW receiver based on a nonmaximally decimated, polyphase realization of the Short Time Fourier Transform (STFT). This patented implementation, (Serial Number 08/816,951), reduces the required encoder processing speed requirements by a factor of 16 while simultaneously allowing processing of multiple, filtered sinusoids (99).
- Construction of a new and simplified ILS algorithm for estimating the frequency of a single complex sinusoid in additive white noise. The method developed exceeds the accuracy of other single sinusoid estimation techniques and can be implemented using simple vector product operations (96).
- Establishment of an approximate complex model of the data at the output of the EW receiver. The derivations show the decimated complex output, actually generated from the sum of real filtered sinusoids in real noise, can be approximately modeled as being generated from the sum of complex filtered sinusoids in complex noise.

The results of this research, whether taken individually or collectively, represent a major contribution to the theory of signal processing and parameter estimation. In particular, this research builds the bridge connecting sinusoidal frequency estimation with LP linear system modeling. Finally, this dissertation derives the connection between real and complex sinusoidal parameter estimation. The estimators constructed as a result of this connection will significantly improve the operational envelope of the Air Force's next generation EW receiver.

### *1.5 Overview*

Chapter II reviews the current literature concerning real sinusoidal parameter estimation techniques. Several estimation algorithms are briefly discussed to show that current

research in the area of parameter estimation has neglected the effects of a filter upon the sinusoids and noise.

Chapter III derives estimators for the amplitude, phase and frequency parameters of filtered sinusoids in white noise. After describing the effects of a filter on sinusoids and noise, a new data model is developed and ML estimators for the sinusoidal parameters are constructed based on this model. Frequency estimation is then recast as the estimation of the coefficients parameterizing the LP general linear model. After developing algorithms for estimating these coefficients, this chapter concludes by introducing a novel method to assess the accuracy of the coefficient and, consequently, the frequency estimates.

Chapter IV applies these estimation algorithms to a specific EW receiver architecture. This chapter begins by deriving the architecture characteristics and showing how substantial processing speed reductions can be obtained by employing multirate signal processing techniques. After modeling both the real and complex forms of the data, the performance of the estimation algorithms, within the operational envelope of the receiver, is then documented and shown to provide improved parameter estimates.

Chapter V summarizes the original contributions of this dissertation.

## II. Literature Review: Sinusoidal Frequency Estimation

### 2.1 Introduction

This chapter reviews the current methods of estimating the frequencies of the sum of  $P$  real sinusoids,  $s[m]$ , in additive noise,  $w[m]$ , from  $M$  observations  $y[m]$ , using the model

$$y[m] = \sum_{k=1}^P b_k \cos(2\pi f_k m + \phi_k) + w[m] = s[m] + w[m] \quad (2.1)$$

Here, for  $0 \leq m \leq M - 1$ , the amplitudes,  $\{b_k\}$ , phases,  $\{\phi_k\}$ , and frequencies,  $\{f_k\}$ , are deterministic but unknown. Since the amplitudes and phases can be found by a linear least squares fit to the data once the frequencies are calculated (45), (71), most research in parameter estimation has focused on developing accurate frequency estimators. This chapter begins by reviewing several methods of estimating the frequency of a single sinusoid and concludes by examining the current methods for  $P$  sinusoids based on linear predictive modeling.

### 2.2 Single Sinusoid Frequency Estimation

In many signal processing applications, the time and frequency components of a block of data samples can be limited so that only one sinusoid exists throughout the data block (71). For example, a bandpass filter can be used to limit the frequency range under observation such that the filtered output can be considered as being from a single sinusoid (90:21).

One method of estimating the frequency of a single sinusoid is via an amplitude search of a set of Discrete Fourier Transform (DFT) coefficients followed by a frequency interpolation based on these coefficients (71)

$$\hat{f}_1 = \frac{1}{M} \left[ k + \frac{|Y(k+1)|}{|Y(k+1)| + |Y(k)|} \right] \quad (2.2)$$

Here  $|Y(k+1)|$  and  $|Y(k)|$  represent the two largest adjacent DFT magnitudes. Since this technique produces only coarse estimates which are dependent on the block length, further optimization techniques must be employed to improve estimation accuracy (85).

An alternative method for estimating the frequency of a single sinusoid employs the conversion of the real sinusoid to its complex form. That is, the complex form of a sinusoid in noise can be written as

$$\tilde{u}[m] = b_1 e^{j(\omega_1 m + \phi_1)} + \tilde{w}[m] = \tilde{b}_1 e^{j\omega_1 m} \{1 + \tilde{v}[m]\} \quad (2.3)$$

Here,  $\omega$  is the radian frequency and  $\tilde{v}[m] = v_R[m] + jv_I[m]$  is a complex random noise sequence described in Appendix G. One approach for estimating  $\omega_1$ , and hence  $f_1$ , relies on extracting the instantaneous phase of  $\tilde{u}[m]$ . Since the frequency of an analog signal can be obtained by differentiating its instantaneous phase (5), a similar relationship exists in the digital domain when the instantaneous phase of the complex sinusoids is given by  $\theta_1(m) = \omega_1 m + \phi_1$ . Provided  $\theta_1(m) > \theta_1(m-1)$ , the frequency can be determined via simple phase differencing (88). Assuming  $b_1$  is large relative to the noise sample magnitudes, as shown in Appendix J,  $\tilde{u}[m]$  can be written as (89)

$$\tilde{u}[m] = b_1 e^{j(\omega_1 m + \phi_1 + v_I[m])} = b_1 e^{j\theta(m)} \quad (2.4)$$

Thus, large SNRs allow  $\theta(m)$  to be treated as a linear function of the frequency,  $\omega_1$ , and the phase,  $\phi_1$ , corrupted by additive noise,  $v_I[m]$ .

This linear phase characteristic has been exploited to provide several simple frequency estimators. One method, termed the Weighted Phase Averaging (WPA) algorithm, extracts the phase after summing  $N$  samples of  $\tilde{u}[m]$  to achieve the new sequence  $\tilde{y}[m]$ , (36), (42),

$$\tilde{y}[m] = \frac{1}{N} \sum_{n=0}^{N-1} \tilde{u}[m-n] = \frac{1}{N} \sum_{n=0}^{N-1} b_1 e^{j(\omega_1 [m-n] + \phi_1)} + \tilde{w}[m-n] \quad (2.5)$$

If  $b_1$  is large in comparison to the noise sample magnitudes,  $\tilde{y}[m]$  is approximated as (42)

$$\tilde{y}[m] \approx b_1 e^{j(\omega_1 m + \phi - \psi)} \alpha e^{-j\gamma[m]} = \alpha b_1 e^{j\theta(m)} \quad (2.6)$$

Here, the quantities  $\psi = \omega_1 [N-1]/2$  and  $\alpha = \sin(\omega_1 N/2) [\sin(\omega_1/2)]^{-1}$  are real numbers and the variable  $\gamma[m] = \alpha^{-1} \sum_{n=0}^{N-1} v_I[m-n]$  is random. Thus, for  $m = N-1 \dots M-1$ ,

the resulting phase difference,  $z[m] = \theta[m] - \theta[m-1]$ , has the vector form

$$\mathbf{z} = 2\pi f_1 \mathbf{1} + [N\alpha]^{-1} G \mathbf{v}_I \quad (2.7)$$

Here  $\mathbf{1}$  is a vector of ones while  $G$  is a sparse matrix defined by

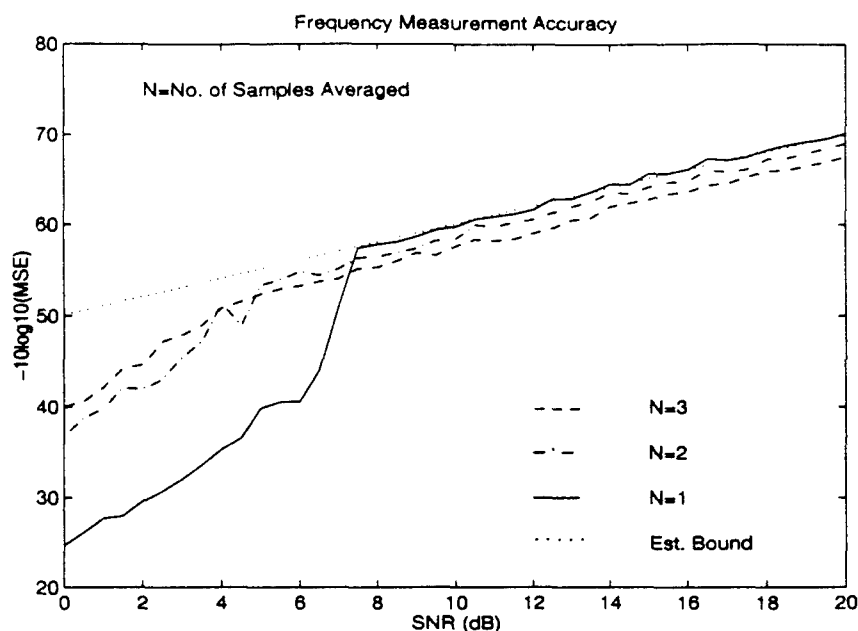
$$[G]_{ij} = \begin{cases} 1 & \text{for } i = j \\ -1 & \text{for } j = i + N \\ 0 & \text{otherwise} \end{cases} \quad (2.8)$$

If  $\tilde{w}[m]$  is a zero mean, normally distributed, independent, complex random variable with variance  $\sigma^2$ , then, as shown in Appendix J,  $\mathbf{z}$  will be a normally distributed random vector with a mean of  $2\pi f_1 \mathbf{1}$  and covariance matrix  $K_{\mathbf{z}} = [N\alpha\sqrt{2}b_1]^{-2} \sigma^2 G G^T$ . As such, an unbiased, minimum variance estimator for  $f_1$  is found via linear regression on  $\mathbf{z}$  as (40:97)

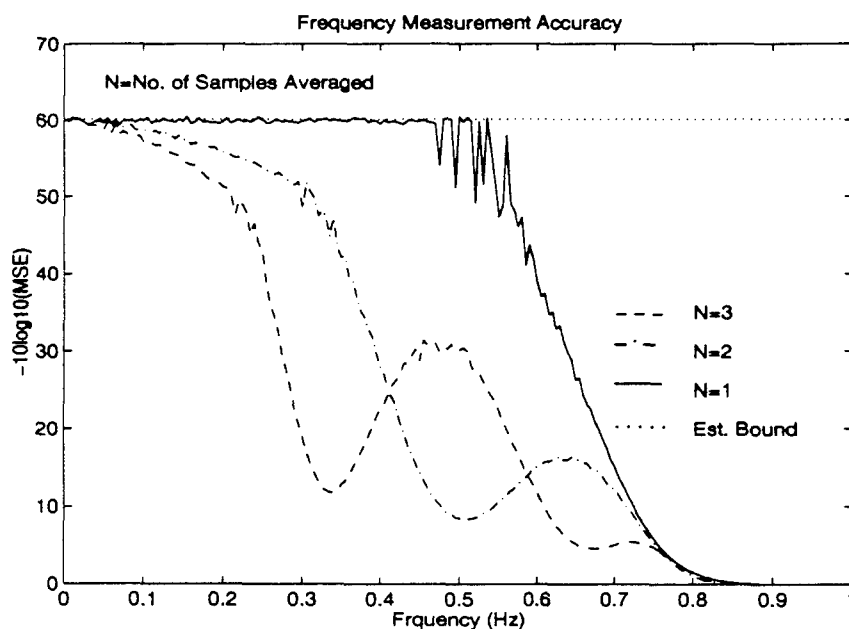
$$\hat{f}_1 = \frac{\mathbf{1}^T [G G^T]^{-1} \mathbf{z}}{2\pi \mathbf{1}^T [G G^T]^{-1} \mathbf{1}} \quad (2.9)$$

Figure 2.1 shows the accuracy of this estimator varies as a function of  $N$ , the SNR, and the frequency of the sinusoid (10). For  $N > 1$ ,  $\tilde{y}[m]$  can be shown to be the output of a low pass filter of length  $N$  (64:447); the improvement in SNR for low frequency sinusoids is translated into a lower SNR threshold. However, once this threshold is reached, the estimator with  $N = 1$  outperformed the other values of  $N$  tested both as a function of SNR and frequency. The reduction in accuracy for  $f > .5$  is due to phase wrapping ambiguities. Provided the SNR is high enough and the frequency range of the input sinusoid is limited, this estimator will provide accurate frequency estimates and is relatively simple to implement for real-time applications (17), (47). In addition, a simple modification of this algorithm allows the rate of change of the frequency to be estimated in addition to the carrier frequency (12), (34).

An alternative method, termed the Iterative Phase Averaging (IPA) algorithm, utilizes an iterative approach for estimating the frequencies from the phase (85). Let  $f(k)$  be the  $k^{th}$  estimate of the frequency and define the variable  $g_k[m] = \tilde{u}[m]e^{-j\omega_k m}$ . Then  $g_k[m]$



(a) Sinusoid Parameters:  $[b_1 = 1, f_1 = .2, \phi = \pi/8]$   
MSE calculated at SNR intervals of .5dB.



(b) Sinusoid Parameters:  $[b_1 = 1, \phi = \pi/8]$ ; SNR = 10dB and  
MSE calculated at frequency intervals of .001 Hz.

Figure 2.1 **WPA Frequency Estimation Accuracy:** MSE calculated from 500 independent realizations of  $M = 25$  samples of sinusoid in noise.  $\text{SNR} = -10\log_{10}\{\sigma^2\}$  and  $\text{MSE} = \frac{1}{500} \sum_{i=1}^{500} (\hat{f}_1[i] - f_1)$ .

satisfies the first order difference equation

$$e^{-j\gamma_k} g_k[m+1] - e^{j\gamma_k} g_k[m] = 0 \quad (2.10)$$

where  $\gamma_k = \frac{1}{2}(\omega_1 - \omega_k)$ . For an arbitrary window function,  $t[m]$ , of finite length  $M$ , then

$$\sum_{m=0}^{M-1} (t[m]e^{-j\gamma_k} - t[m+1]e^{j\gamma_k}) g_k[m] = 0 \quad (2.11)$$

Using the real and imaginary components to find the phase yields the iterative estimator

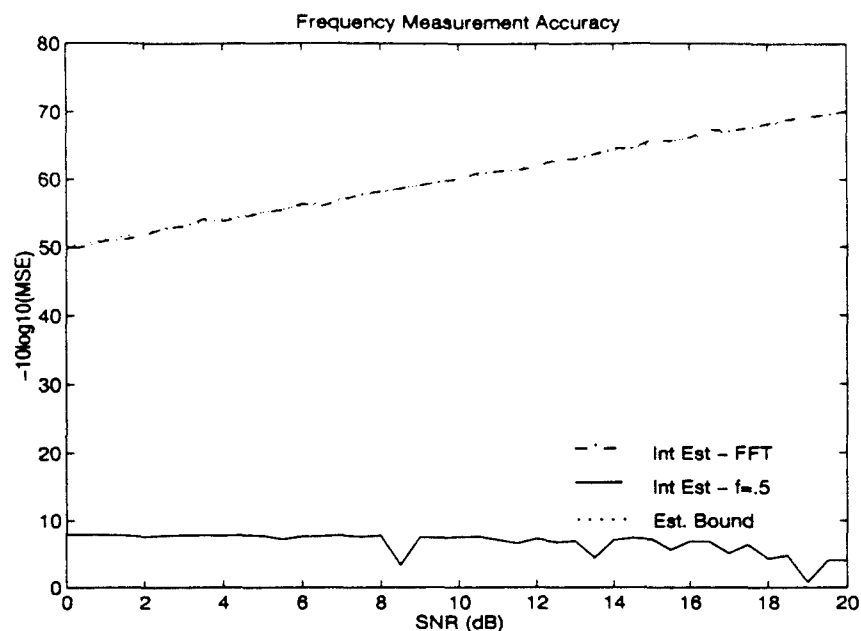
$$\hat{f}(k+1) = \hat{f}(k) + \frac{j \sum_{m=0}^{M-1} (t[m+1] - t[m]) \tilde{u}[m] e^{-j\omega_k m}}{\pi \sum_{m=0}^{M-1} (t[m+1] + t[m]) \tilde{u}[m] e^{-j\omega_k m}} \quad (2.12)$$

For a complex sinusoid, the optimal window is the parabola given by  $t[m] = m(M-m)$  (85). The main problem involved with using this estimator is obtaining a good initial frequency estimate. Typically, the initial estimate is obtained from the Fast Fourier Transform (FFT) bin with the maximum magnitude (85). As shown in Figure 2.2 this algorithm provides accurate frequency estimates, both as a function of SNR and frequency, only when a good initial estimate is attained. In addition, provided the noise can be modeled as a zero mean, complex, normally distributed, uncorrelated random variable, this algorithm will provide unbiased minimum variance frequency estimates when provided with a good initial estimate (85).

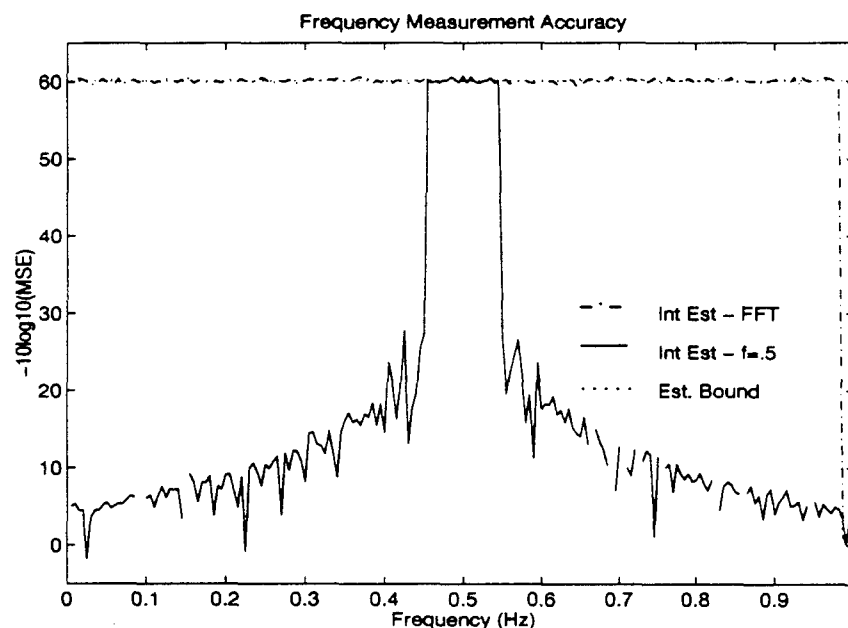
### 2.3 Multiple Sinusoid Frequency Estimation

In general, for  $P$  real sinusoids in noise, accurate frequency estimation can be difficult. If the sinusoids are well separated in frequency, the DFT can be used to provide accurate frequency estimates. However, if the sinusoids are not approximately orthogonal over the observation interval, DFT processing is ineffective and other techniques must be employed (45). For these cases, direct maximum likelihood techniques have been shown to provide accurate frequency estimates (82). These techniques usually require an iterative, nonlinear multidimensional search in the frequency domain and are too computationally intense to be implemented in real-time.





(a) Sinusoid Parameters:  $[b_1 = 1, f_1 = .2, \phi = \pi/8]$   
MSE calculated at SNR intervals of .5dB.



(b) Sinusoid Parameters:  $[b_1 = 1, \phi = \pi/8]$ ; SNR = 10dB and  
MSE calculated at frequency intervals of .001 Hz.

Figure 2.2 **IPA Frequency Estimation Accuracy:** MSE calculated from 500 independent realizations of  $M = 25$  samples of sinusoid in noise.  $\text{SNR} = -10 \log_{10}\{\sigma^2\}$  and  $\text{MSE} = \frac{1}{500} \sum_{i=1}^{500} (\hat{f}_1[i] - f_1)$ .

Alternatively, a simpler method for determining the frequencies can be accomplished using linear predictive modeling techniques (57), (95). Let  $s[m]$  be a real signal consisting of the sum of  $P$  real sinusoids, or modes. A  $2P^{th}$  order polynomial,  $A(z)$ , which incorporates these modes can be constructed as (73:484-485)

$$A_o(z) = a[0] \prod_{k=1}^P (1 - z_k z^{-1})(1 - z_k^* z^{-1}) = \sum_{p=0}^{2P} a[p] z^{-p} \quad (2.13)$$

For  $P$  real sinusoids, the LP coefficients,  $a[p]$ , are real and symmetric with  $a[p] = a[2P-p]$ ; the  $P$  frequencies present in  $s[m]$  are roots to  $A(z)$  (8). In the time domain, the polynomial becomes the Linear Prediction (LP) equation

$$\sum_{p=0}^{2P} a[p] s[m-p] = 0 \quad (2.14)$$

When the sinusoids are corrupted with noise, the LP equation is satisfied only in a statistical sense. Consequently, several algorithms based on algebraic, iterative filtering, and maximum likelihood techniques, have been developed to estimate the LP coefficients from the observations,  $y[m]$ .

**2.3.1 Algebraic Techniques.** One technique of identifying the LP coefficients uses an extended LP model of length  $L > 2P$  to account for noise modes (37) and constrains the LP coefficients to satisfy the Forward-Backward LP (FBLP) requirements (45). The resulting system of equations, formed by the substitution of  $y[m]$  for  $s[m]$  in the LP equation, is then solved using algebraic techniques to obtain the  $L^{th}$  order polynomial  $A(z)$ . The  $P$  roots of  $A(z)$ , whose magnitudes are closest to the unit circle, for  $0 < f < .5$ , are then used to provide the frequency estimates (43).

To construct the system of equations, define  $\mathbf{a}$  as the vector of  $L+1$  linear prediction coefficients,  $\mathbf{a} = [1, \mathbf{a}_o]^T$ , where the vector  $\mathbf{a}_o$  consists of the truncated form of  $\mathbf{a}$  and defined as  $\mathbf{a}_o = [a[1] \dots a[L]]^T$ . In addition, define the  $M-L$  by  $L+1$  forward observation matrix,  $Y_F$ , and the backward observation matrix,  $Y_B$ , as

$$[Y_F]_{k,l} = y[M+1-k-l]; \quad [Y_B]_{k,l} = y[M-L-1-k+l] \quad (2.15)$$

for  $k = 1 \dots M - L$  and  $l = 1 \dots L + 1$ . Concatenating  $Y_F$  and  $Y_B$  into a single matrix gives the FB observation matrix  $Y_{FB}$  (45)

$$Y_{FB}^T = [Y_F^T Y_B^T] = [\mathbf{y}_{fb} | Y_{fb}]^T \quad (2.16)$$

where  $\mathbf{y}_{fb}$  is the first column of  $Y_{FB}$ . Finally, define the FB signal matrix,  $S_{FB}$ , in a manner similar to  $Y_{FB}$ . The resulting system of equations to be solved becomes

$$Y_{fb} \mathbf{a}_o = -\mathbf{y}_{fb} \quad (2.17)$$

One algebraic technique for solving this equation is based on a Least Squares (LS) solution. In general, the optimal LS solution requires the rank  $2P$  signal model matrix,  $S_{fb}$ , to be known exactly (43). Since the components  $s[m]$  are not available, the method proposed by Prony simply substitutes  $y[m]$  for  $s[m]$  to obtain the LS estimate (73:406-408,491-493)

$$\hat{\mathbf{a}}_o = -[Y_{fb}^T Y_{fb}]^{-1} Y_{fb}^T \mathbf{y}_{fb} \quad (2.18)$$

For a fixed number of measurements, this estimator will provide accurate estimates only for large SNRs (82). In addition, as the number of observations,  $M$ , approaches infinity, this estimator can be shown to be asymptotically biased (14).

Figure 2.3 shows the estimation accuracy of the LS algorithm for two real sinusoids in zero mean, independent, normally distributed noise. For these sinusoids, above an SNR of about 15dB the estimation accuracy improved as  $L$  increased from 4 to 14. This indicates the extra coefficients are adequately modeling noise. However, this improvement in estimation accuracy comes at the expense of rooting a large-order polynomial.

An alternate algebraic technique for obtaining the LP coefficients from Equation 2.17 is known as the Principal Eigenvalue (PE) method (44), (45), (91). This method provides an estimate of the LP coefficients by first performing an eigenvalue analysis of the matrix product  $[Y_{fb}^T Y_{fb}]$  and then providing the estimate of  $\hat{\mathbf{a}}_o$  based on this analysis (45), (91).

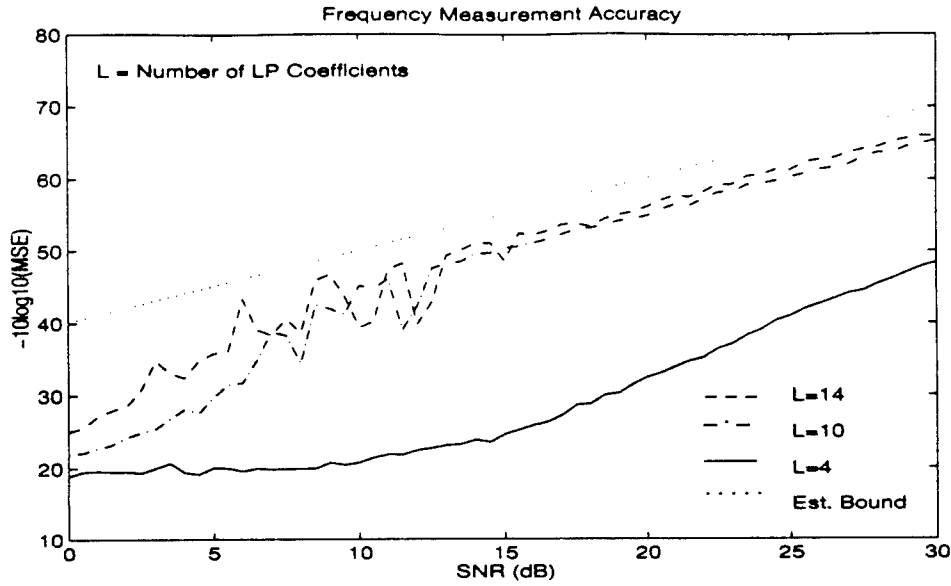


Figure 2.3 **LS Frequency Estimation Accuracy:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 30$  samples of two sinusoids in noise. Sinusoid Parameters:  $[b_1 = 1, f_1 = .2, \phi_1 = \pi/8]$ ,  $[b_2 = 1, f_2 = .22, \phi_2 = \pi/3]$ ;  $SNR = -10 \log_{10} 2\sigma^2$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ ; Frequencies chosen from two roots of  $A(z)$  with magnitudes closest to unit circle for  $0 < f < .5$ .

The resulting estimate of the LP coefficients becomes

$$\hat{\mathbf{a}}_o = \sum_{l=1}^{2P} \lambda_l^{-1} \mathbf{v}_l \mathbf{v}_l^T \left( -Y_{fb}^T \mathbf{y}_{fb} \right) \quad (2.19)$$

Here, the vectors  $\mathbf{v}_l$  are the orthonormal eigenvectors of the matrix product  $[Y_{fb}^T Y_{fb}]$  and  $\lambda_l$  is an associated eigenvalue with  $\lambda_1 \geq \lambda_2 \dots \geq \lambda_{2P}$  and  $\{\lambda_l\} \approx 0$  for  $l = 2P + 1 \dots L$ . The effect of using this truncated representation is to increase the SNR in the data prior to solving for  $\hat{\mathbf{a}}_o$  (44).

Figure 2.4 shows the estimation accuracy of the PE algorithm for two real sinusoids in zero mean, independent, normally distributed noise. For these sinusoids, above an SNR of about 8dB the estimation accuracy improved as  $L$  increased from 4 to 12. This indicates the extra coefficients are adequately modeling noise. However, this improvement in estimation accuracy comes at the expense of rooting a large order polynomial. In addition, the improvement in the SNR threshold over the LS algorithm is due to the use of only the four largest eigenvalues in the estimate of  $\hat{\mathbf{a}}_o$ .

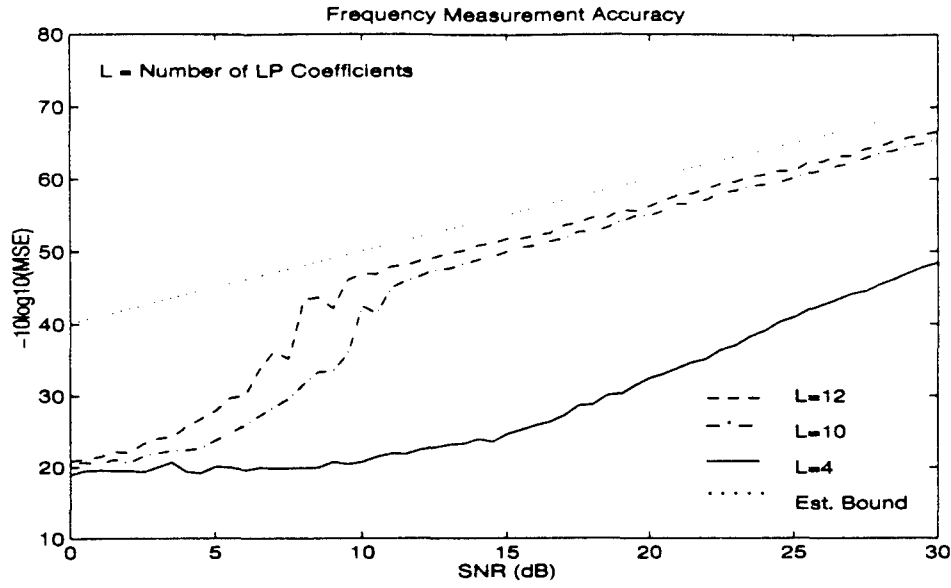


Figure 2.4 **PE Frequency Estimation Accuracy:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 30$  samples of two sinusoids in noise. Sinusoid Parameters:  $[b_1 = 1, f_1 = .2, \phi_1 = \pi/8]$ ,  $[b_2 = 1, f_2 = .22, \phi_2 = \pi/3]$ ;  $SNR = -10 \log_{10} 2\sigma^2$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ ; Frequencies chosen from two roots of  $A(z)$  with magnitudes closest to unit circle for  $0 < f < .5$ .

A final algebraic technique of estimating the LP coefficients is based on a Total Least Squares (TLS) solution to Equation 2.17(67), (30). That is, with  $a[0] = 1$ , the system of equations to be solved becomes

$$Y_{FB} \mathbf{a} = 0 \quad (2.20)$$

Since  $y[m] = s[m] + w[m]$ , the signal matrix,  $S_{fb}$ , is now corrupted with noise and can be represented as  $Y_{fb} = S_{fb} + E_{fb}$  where  $E_{fb}$  can be interpreted as the matrix of errors (21). The TLS solution for this over-determined set of equations is based on a Singular Valued Decomposition (SVD) of  $Y_{FB}$ , given by  $Y_{FB} = U \Sigma V^T$ , so that the estimate for  $\mathbf{a}$  becomes

$$\hat{\mathbf{a}} = \begin{bmatrix} 1 \\ \hat{\mathbf{a}}_o \end{bmatrix} = \frac{\sum_{l=2P+1}^{L+1} V(1, l) \mathbf{v}_l}{\sum_{l=2P+1}^{L+1} V(1, l) V(1, l)} \quad (2.21)$$

where  $\mathbf{v}_l$  is the  $l^{th}$  column of  $V$ , the matrix containing the right singular vectors of  $Y_{FB}$  and  $V(1, l)$  is the first element of the  $l^{th}$  column of  $V$ .

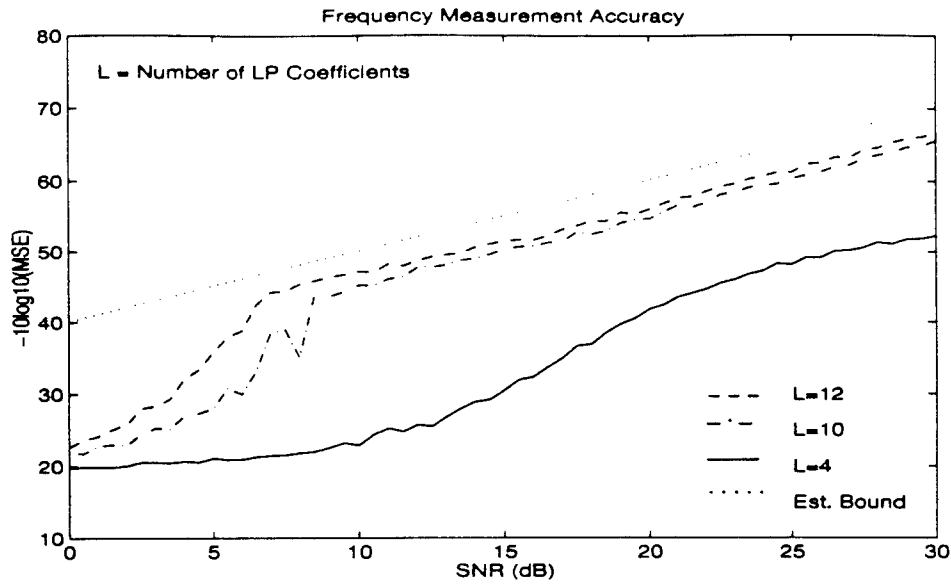


Figure 2.5 **TLS Frequency Estimation Accuracy:** MSE calculated at SNR intervals .5dB from 500 independent realizations of  $M = 30$  samples of two sinusoids in noise. Sinusoid Parameters:  $[b_1 = 1, f_1 = .2, \phi_1 = \pi/8]$ ,  $[b_2 = 1, f_2 = .22, \phi_2 = \pi/3]$ ;  $SNR = -10 \log_{10} 2\sigma^2$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ ; Frequencies chosen from two roots of  $A(z)$  with magnitudes closest to unit circle for  $0 < f < .5$ .

Figure 2.5 shows the estimation accuracy of the TLS algorithm for two real sinusoids in zero mean, independent, normally distributed noise. For these sinusoids, above an SNR of about 6dB the estimation accuracy improved as  $L$  increased from 4 to 12. Again, this indicates the extra coefficients are adequately modeling noise. As with the LS and PE estimates, this improvement in estimation accuracy comes at the expense of rooting a large-order polynomial. In addition, the TLS has a lower SNR threshold than either the LS or the PE algorithm. This is due to the fact the TLS algorithm accounts for the noise in the observation matrix as well as the noise in the observation vector (67).

**2.3.2 Iterative Filtering.** Another technique for identifying the LP coefficients exploits the property that any regular stationary random process can be represented as the output of a linear system driven by white noise. This allows estimation of the LP coefficients to be accomplished using AR modeling techniques (59), (87:411-412). A popular method for sinusoids is based on the Steiglitz-McBride method of linear system identification via iterative filtering (78), (81), (82). The basic technique is to iteratively filter

the observations,  $y[m]$ , with the inverse filter,  $1/A(\rho z)$ , derived from the LP coefficient estimates, to obtain a new sequence  $y_d[m]$ . Here,  $\rho$  is a convergence parameter,  $0 < \rho < 1$ , used to ensure stability of the inverse filter (14), (69). New estimates of the LP coefficients are then obtained from analysis of this filtered data. Ideally, since the roots of  $A(z)$  correspond to the frequencies of the sinusoids present in  $y[m]$ , the frequency response of the inverse filter  $1/A(\rho z)$  can be interpreted as  $2P$  narrow bandpass filters centered near these frequencies(69). Since the sinusoidal component of the filtered output,  $y_d[m]$ , consists of sinusoids at the same frequencies as the input sinusoids, the effect of the iterative filtering is to increase the SNR at each iteration. The two methods most often used for LP coefficient identification are the Iterative Filtering Algorithm (IFA) and Parametric Filtering (PF) Algorithm.

The IFA estimator exploits the LP orthogonality principle to find the coefficient vector  $\mathbf{a}$  (87:339). Let  $y[m]$  be estimated from  $2P$  previous samples as

$$\hat{y}[m] = \sum_{p=1}^{2P} a[p]y[m-p] \quad (2.22)$$

and let  $e[m]$  be the error between the estimate and the actual value of  $y[m]$ . Assuming  $a[0] = 1$ , the coefficient vector,  $\mathbf{a}$ , which minimizes the mean square error,  $E\{e[m]e[m]\}$ , is found such that the error is orthogonal to the previous observations (87:338-341)

$$E\{e[m]y[m-k]\} = 0 = \sum_{l=0}^L a[l]r_{yy}[k-l] \quad (2.23)$$

for  $k = 1 \dots 2P$  and  $r_{yy}[k-l] = E\{y[k]y[l]\}$ . Since  $r_{yy}[l]$  is not known, it too must be estimated from the data (87:514-577). The IFA estimator uses the Burg algorithm with  $\rho = 1$  to estimate  $r_{yy}[l]$  and the LP coefficients while ensuring the resulting polynomial,  $A(z)$ , is stable(39:161-171,228-229). In addition, an appropriate nonnegative window, such as the Hamming or Optimal Tapered Burg Window, is often employed to reduce the dependency of this estimator upon the phase of the sinusoids, (13), (35), (39:231). The LP coefficients are then iteratively estimated using the filtered data and the algorithm terminates upon convergence of the LP estimates (38), (39:417-419). This estimator is

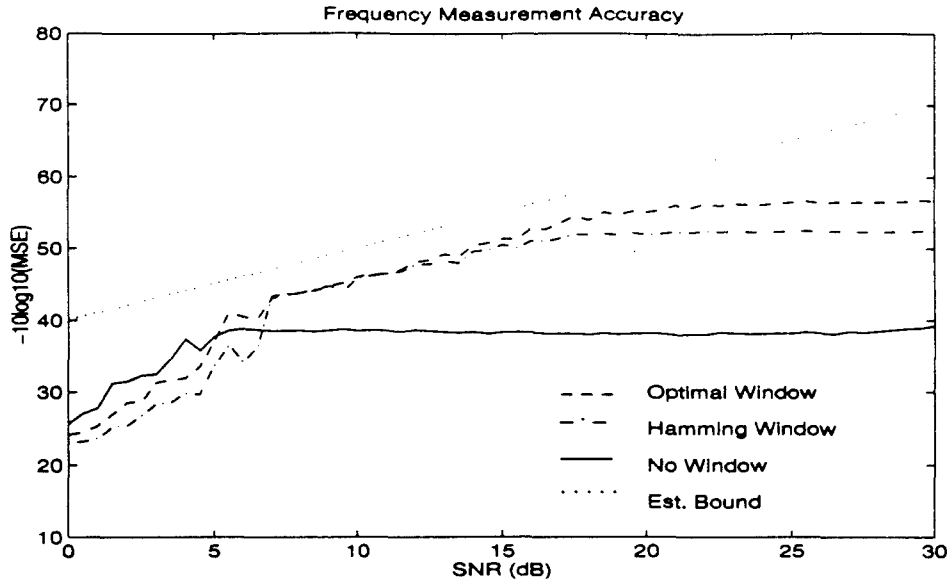


Figure 2.6 **IFA Frequency Estimation Accuracy:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 30$  samples of two sinusoids in noise. Parameters:  $[b_1 = 1, f_1 = .2, \phi_1 = \pi/8]$ ,  $[b_2 = 1, f_2 = .22, \phi_2 = \pi/3]$ ;  $SNR = -10\log_{10} 2\sigma^2$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (f_p(i) - f_p)^2$ .

relatively simple to implement and can be shown to provide maximum likelihood frequency estimates for a large number of measurements (39:417-419).

Figure 2.6 shows the estimation accuracy of the IFA estimator for two real sinusoids in zero-mean, independent, normally distributed noise. For these sinusoids, above an SNR of about 7dB, use of the Optimal Tapered Burg Window provides the most accurate frequency estimates over the SNR range of 7dB to 20dB. Regardless of the window used however, the flattening of the MSE curves indicates this estimator becomes biased at high SNRs for this small number measurements.

As an alternative to the IFA correlation based approach, the PF algorithm institutes the sinusoidal constraints,  $a[p] = a[2P - p]$ , and uses an LS methodology for estimating the LP coefficients(14), (50), (51), (52), (53), (94). Here, the original observations,  $y[m]$ , are passed through a causal, infinite impulse response filter which is parameterized by a set of symmetric scaling coefficients,  $d[p]$ , to obtain the filtered data  $y_d[m]$  as

$$y_d[m] = - \sum_{p=1} d[p] \rho^p \hat{a}[p] y_d[m-p] + y[m] \quad (2.24)$$



For  $P$  real sinusoids, the scaling coefficients,  $\{d[p]\}$ , are calculated as, (53).

$$d[q] = \frac{1 + \rho^{2q}}{\rho^q + \rho^{2P-q}} \quad (2.25)$$

for  $q = 1 \dots P$ . To obtain the LS estimate of the LP coefficients, the PF algorithm imposes the sinusoidal constraints on  $\mathbf{a}$  via a  $2P + 1$  by  $P$  matrix  $B$  so that  $\mathbf{a} = B\boldsymbol{\alpha}$ . Here, the vector  $\boldsymbol{\alpha}$  is given by  $\boldsymbol{\alpha} = [1, a[1] \dots a[P]]^T$  is the set of constrained LP coefficients. Defining  $Y_C$  as the constrained observation matrix,  $Y_C = Y_F(\mathbf{d})B = [\mathbf{y}_c | Y_c]$ , where  $\mathbf{y}_c$  is the first column of  $Y_C$ , yields the PF estimator

$$\hat{\boldsymbol{\alpha}}_o = -[Y_c^T Y_c]^{-1} Y_c^T \mathbf{y}_c \quad (2.26)$$

The LP coefficients are then iteratively estimated using the filtered data and the algorithm terminates upon convergence of the LP estimates. This estimator is relatively simple to implement and can be shown to be asymptotically unbiased as  $M$  approaches infinity (50).

Figure 2.7 shows the estimation accuracy of the PF algorithm for two real sinusoids in zero mean, independent, normally distributed noise. For these sinusoids, above an SNR of about 4dB, use of the convergence parameter of  $\rho = .96$  provides the most accurate frequency estimates over the SNR range of 4dB to 15dB. However, the flattening of the MSE curves indicates this estimator becomes biased at high SNRs when the number of data points is finite.

**2.3.3 Maximum Likelihood.** The final technique considered for estimating the LP coefficients requires minimization of the following objective function,  $J(\mathbf{a})$ , with respect to the LP coefficients in  $\mathbf{a}$  (6), (46), (68)

$$J(\mathbf{a}) = \mathbf{a}^T Y_F^T (A^T A)^{-1} Y_F \mathbf{a} \quad (2.27)$$

Here,  $\mathbf{a} = [a[0] \dots a[2P]]^T$  while  $A$  is the  $M$  by  $M - 2P$  matrix of LP coefficients

$$[A]_{k,l} = \begin{cases} a[k-l] & \text{for } l \leq k \leq l+2P \\ 0 & \text{otherwise} \end{cases} \quad (2.28)$$

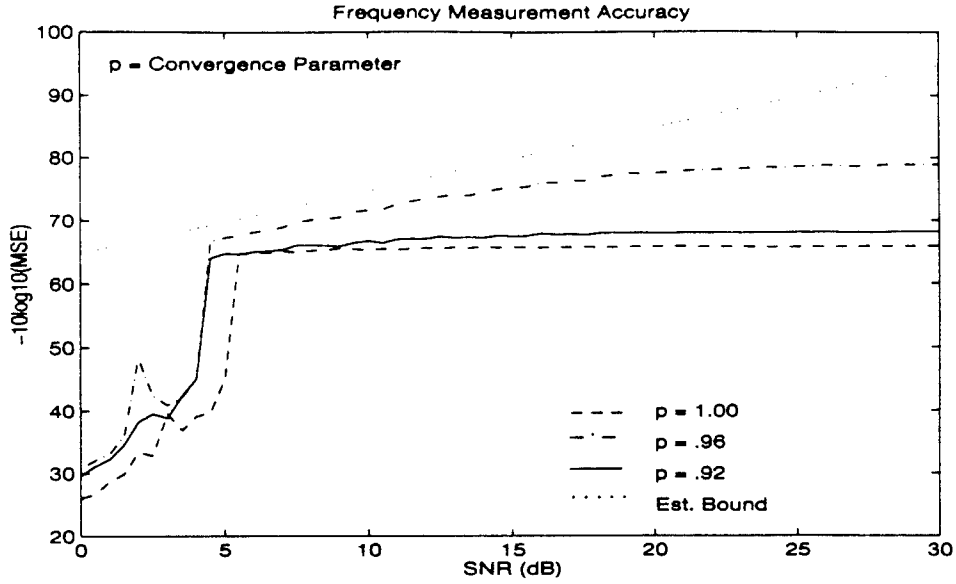


Figure 2.7 **PF Frequency Estimation Accuracy:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 100$  samples of two sinusoids in noise. Parameters:  $[b_1 = 1, f_1 = .2, \phi_1 = \pi/8]$ ,  $[b_2 = 1, f_2 = .22, \phi_2 = \pi/3]$ ;  $SNR = -10 \log_{10} 2\sigma^2$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ .

The frequencies found from the LP coefficients which minimize  $J(a)$  provide maximum likelihood frequency estimates (6), (63). Letting  $\alpha = [a[0] \dots a[P]]^T$  and imposing the sinusoidal constraints via the  $2P + 1$  by  $P + 1$  matrix  $B$  yields the new objective function

$$J(\alpha) = \alpha^T B^T Y_F^T (A^T A)^{-1} Y_F B \alpha \quad (2.29)$$

The most popular method of minimizing this equation is termed the Iterative Quadratic Maximum Likelihood (IQML) algorithm (6), (46), (68). This algorithm constructs the matrix  $A$  from the current estimate of  $\alpha$  and minimizes  $J(\alpha)$  with respect to  $\alpha$  by imposing the constraint that  $\alpha^T \alpha = 1$  (75). Termination of this algorithm is reached upon convergence of  $\alpha$ . Though this algorithm involves an eigenvalue analysis of a  $P + 1$  by  $P + 1$  square matrix at each iteration, relatively efficient methods of algorithm implementation have been developed (9), (29), (46). Indeed, this algorithm is currently the most popular method of estimating the frequencies of multiple complex sinusoids in zero-mean, independent, normally distributed noise (6), (46), (68), (63).

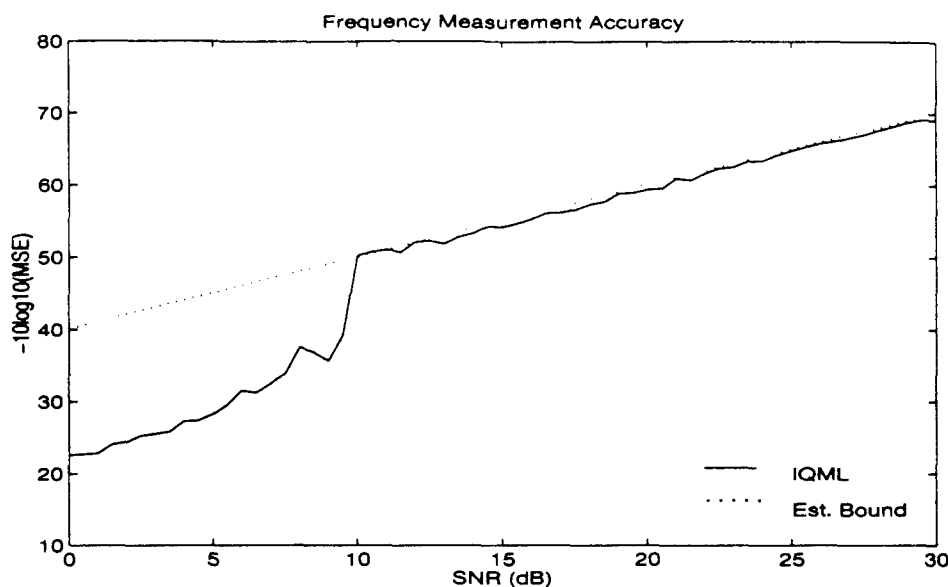


Figure 2.8 **IQML Frequency Estimation Accuracy:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 30$  samples of two sinusoids in noise. Sinusoid Parameters:  $[b_1 = 1, f_1 = .2, \phi_1 = \pi/8]$ ,  $[b_2 = 1, f_2 = .22, \phi_2 = \pi/3]$ ;  $SNR = -10 \log_{10} 2\sigma^2$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ .

Figure 2.8 shows the estimation accuracy of the IQML algorithm for two real sinusoids in zero-mean, independent, normally distributed noise. For these sinusoids, accurate frequency estimates were attained for SNRs above about 10dB using an initial estimate of  $\alpha = [1, 0 \dots 0]^T$ . An improvement in the SNR threshold should be possible if a better initial estimate of  $\alpha$  were used. In addition, the fact that optimal estimates were attained for all SNRs above the threshold indicates this algorithm can be used as a general frequency estimator for large SNRs.

#### 2.4 Summary

This chapter reviewed several frequency estimators for  $P$  real sinusoids in noise. For  $P = 1$ , relatively simple estimators can be constructed by extracting the phase of the complex representation of the sinusoid. Simulations show these estimators will achieve optimal frequency estimates provided the noise is a zero-mean, independent, normally distributed, complex random variable.

For  $P > 1$ , estimation of the frequencies was recast as the estimation of a set of LP coefficients and several methods for estimating the coefficients were examined. The first technique utilized the FBLP constraints to construct a linear system of equations based on an  $L^{\text{th}}$  degree polynomial. Simulations show that above a given SNR threshold, the resulting LS, PE, and TLS estimators can provide relatively accurate estimates of the frequencies of sinusoids in zero-mean, normally distributed, uncorrelated noise. This accuracy comes at the expense of determining the frequencies from a large model order. In addition, the optimal length of the model depends on both the number of sinusoids to be estimated and their frequencies. These characteristics prevent this technique from being used as a general estimator for the frequencies of  $P$  real sinusoids in noise.

The second technique employed Iterative Filtering methods to repeatedly filter the data prior to estimation of the LP coefficients. The LP coefficients for a minimum order polynomial were then estimated using the IFA and PF methods. Simulations show that over a range of SNRs, these estimators provide accurate frequency estimates for sinusoids in zero-mean, normally distributed, uncorrelated noise. However, these accuracies are dependent on algorithm parameters and previous research indicates these algorithms can be sensitive to the phases of the sinusoids (13), (35). These limitations prevent this technique from being used as a general estimator for the frequencies of  $P$  real sinusoids in noise.

The third technique used to estimate the LP coefficients was based on maximum likelihood techniques. By recasting maximum likelihood frequency estimation as a minimization of a nonlinear function,  $J(\mathbf{a})$ , with respect to the LP coefficients, maximum likelihood frequency estimates could be obtained as roots of a minimum order LP polynomial. The LP coefficients were then estimated using the IQML algorithm to minimize  $J(\mathbf{a})$ . Simulations show this estimator provides accurate frequency estimates, above a given SNR threshold, for sinusoids in zero-mean, normally distributed, uncorrelated noise. This accuracy comes at the expense of an iterative eigenvalue analysis of a  $P + 1$  by  $P + 1$  square matrix.

This review of the current methods of frequency estimation show that the true linear model relating the LP coefficients to the frequencies of  $P$  real sinusoids in noise has not been adequately established. This lack of a true model is evidenced by the many methods

available for estimating the LP coefficients from the data. If the true model existed, the accuracy of the estimators would be evaluated against the model itself.

In addition, no attempt has been made incorporate the statistical characteristics of the noise into the construction of the LP coefficient estimators. Each of the estimators examined in this chapter were evaluated under the assumption the noise was uncorrelated. However, in general, when analyzing analog signals using digital methods, an antialiasing filter is used to limit the frequency range of the analog signals under consideration. Any noise passing through the filter will be correlated by the action of the filter. Accurate estimation of the LP coefficients will require knowledge of how the noise is correlated.

Finally, there is currently no method to gauge the accuracy of a point estimate of the LP coefficients, and consequently the frequencies, based strictly on a single realization of the data. Knowledge that a point estimate is unbiased is inadequate for making decisions based on the estimation accuracy. A bound on the estimation error must also be reported.

This dissertation will employ stochastic modeling techniques to construct the true linear model relating the LP coefficients to the frequencies of  $P$  real, filtered sinusoids in noise. Once this model is constructed, estimators for the LP coefficients will be derived and their estimation accuracies evaluated. A method for bounding the coefficient and frequency estimation error, based on a single realization, will then be developed and used to establish confidence intervals for the frequency estimates. The estimators developed from this research will then be applied to the Air Force's next-generation EW receiver and evaluated within the receiver's operational envelope.

### *III. Theory: Sinusoidal Parameter Estimation*

#### *3.1 Introduction*

This chapter derives estimators for the amplitude, phase and frequency parameters of filtered sinusoids in white noise. After carefully deriving a mathematical model describing the effects of a filter on sinusoids and noise, Maximum Likelihood (ML) estimators for the sinusoidal parameters are developed based on this model. Estimation of the sinusoidal frequencies is then recast as the estimation of the coefficients of a Linear Prediction (LP) general linear model and algorithms based on fixed point theory are developed for estimating these coefficients. In this dissertation, the relationship between this general linear model and ML frequency estimation is established. Moreover, two popular ML-type methods for estimating the LP coefficients are analyzed in depth and are shown to be fixed point estimators. This chapter concludes by deriving a method to estimate the variance of a point estimate of the LP coefficients and, subsequently, the frequencies, based solely upon a single realization of the measurements and knowledge of the noise variance.

#### *3.2 Filtered Data Model*

This section derives a mathematical model for a real sinusoid passing through a linear, time-invariant filter. With few exceptions, most digital signal processing systems operate on signals which have been already been passed through various systems. In many cases, each system can be modeled as a linear time-invariant filter. The model derived in this section shows how to characterize and account for the effect of a system filter upon the parameter estimation process and is an original contribution of this research.

*3.2.1 Input Signal.* Let the sampled signal,  $x[m]$ , be the sum of a sinusoidal signal,  $\nu[m]$ , and zero-mean, independent, normally distributed noise,  $\eta[m]$ , so that

$$x[m] = \nu[m] + \eta[m] \tag{3.1}$$

Here,  $\nu[m]$  is defined as the linear sum of  $P$  discrete-time real sinusoids

$$\nu[m] = \sum_{k=1}^P b_k^o \cos(2\pi f_k m + \phi_k^o) = \sum_{k=1}^P \nu_k[m] \quad (3.2)$$

where the amplitudes,  $\{b_k^o\}$ , phases,  $\{\phi_k^o\}$ , and frequencies,  $\{f_k\}$ , are assumed to be deterministic but unknown quantities with constraints  $b_k^o > 0$ ,  $0 \leq \phi_k^o < 2\pi$ , and  $0 < f_k < 0.5$ . Assuming there are  $M_o$  samples over which Equation 3.1 holds, in vector form

$$\mathbf{x} = \boldsymbol{\nu} + \boldsymbol{\eta} \quad (3.3)$$

where  $\mathbf{x} = [x[M_o - 1] \dots x[0]]^T$ ,  $\boldsymbol{\nu} = [\nu[M_o - 1] \dots \nu[0]]^T$  and  $\boldsymbol{\eta} = [\eta[M_o - 1] \dots \eta[0]]^T$  are  $M_o$ -dimensional vectors with real components. The problem is to estimate the amplitudes, phases, and frequencies of the sinusoids given the observations,  $\mathbf{x}$ , and knowledge that the noise vector,  $\boldsymbol{\eta}$ , is a zero-mean, normally distributed, random vector with covariance  $K_{\boldsymbol{\eta}} = \sigma^2 I$  where  $I$  denotes the identity matrix. In general, this simple model can not be applied directly to most practical systems since the estimation process is typically accomplished after filtering. Instead, the observations are taken at the output of a known linear system or filter which induces a known covariance structure on the noise.

**3.2.2 Output Signal.** Assume the linear system/filter model can be represented as a finite impulse response, linear, time-invariant filter with  $N$  real coefficients,  $\{h[n]\}$  for  $n = 0 \dots N - 1$ . The frequency transfer function,  $H(e^{j\omega})$ , as function of radian frequency,  $\omega$ , is given by (66:193)

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h[n] e^{-j\omega n} = |H(e^{j\omega})| e^{j\Phi(\omega)} \quad (3.4)$$

where  $\omega = 2\pi f$ . With input  $x[m]$ , for  $N - 1 \leq m \leq M_o - 1$ , the steady state output  $y[m]$  will be real and can be represented as the sum of the filtered sinusoid,  $s[m]$ , and filtered noise,  $w[m]$  (66:192-205). That is

$$y[m] = s[m] + w[m] = \sum_{n=0}^{N-1} h[n] \nu[m - n] + \sum_{n=0}^{N-1} h[n] \eta[n - m] \quad (3.5)$$

Alternatively, since  $y[m]$  is defined to be in the steady state, the sinusoidal component of  $y[m]$ , denoted  $s[m]$ , becomes

$$s[m] = \sum_{k=1}^P b_k \cos(2\pi f_k m + \phi_k) \quad (3.6)$$

Here  $b_k = |H(e^{j\omega_k})|b_k^o$  and  $\phi_k = \Phi(\omega_k) + \phi_k^o$  define the deterministic effects of the filter on the input sinusoidal amplitudes and phases. Now assume there are  $M$  samples of the steady state output with  $M = M_o - N + 1 > 0$ , and define the  $M_o$  by  $M$  Toeplitz filter matrix  $H$  as (40:570)

$$H^T = \begin{bmatrix} h[0] & \dots & h[N-1] & 0 & 0 & \dots & 0 \\ 0 & h[0] & \dots & h[N-1] & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & \vdots & 0 \\ 0 & 0 & \dots & 0 & h[0] & \dots & h[N-1] \end{bmatrix} \quad (3.7)$$

For  $m = N - 1 \dots M_o - 1$ , the output,  $y[m]$ , can be written in vector form as

$$\mathbf{y} = H^T \mathbf{x} = H^T \boldsymbol{\nu} + H^T \boldsymbol{\eta} = \mathbf{s} + \mathbf{w} \quad (3.8)$$

where  $\mathbf{y}$  is defined as an  $M$ -dimensional vector with real components

$$\mathbf{y} = [y[M_o - 1] \dots y[N - 1]]^T$$

while  $\mathbf{s}$  and  $\mathbf{w}$  are defined in a manner similar to  $\mathbf{y}$ . The problem now is to estimate the amplitudes,  $\{b_k\}$ , phases,  $\{\phi_k\}$ , and frequencies,  $\{f_k\}$ , of the sinusoids given the observations,  $\mathbf{y}$ . To accomplish this estimation, a model for the observations will be developed in terms of the frequencies present in the signal and a set of scaling coefficients which are directly related to the amplitudes and phases.

**3.2.3 Model Development.** Since  $M_o > M$ , the columns of  $H$  will be linearly independent, and the noise vector,  $\mathbf{w}$ , becomes a zero-mean, normally distributed, correlated noise sequence with covariance matrix  $K_w$  given by  $K_w = \sigma^2 H^T H$  (86:56). This



$M$  by  $M$  square matrix is assumed to be nonsingular so that  $K_w^{-1}$  exists. The observation vector,  $y$ , can now be described as a normally distributed random vector with a mean vector  $m_y = s$  and a covariance matrix  $K_y = K_w$ . Using trigonometric identities (100:457), Equation 3.5 can be rewritten as

$$y[m] = s[m] + w[m] = \sum_{k=1}^P b_k^c \cos(m\omega_k) - b_k^s \sin(m\omega_k) + w[m] \quad (3.9)$$

where the scaling coefficients,  $b_k^s = b_k \sin(\phi_k)$  and  $b_k^c = b_k \cos(\phi_k)$ , are directly related to the phases and amplitudes (40:198). That is

$$\phi_k = \begin{cases} \tan^{-1} |b_k^s/b_k^c| & \text{if } b_k^s > 0; b_k^c > 0 \\ \pi - \tan^{-1} |b_k^s/b_k^c| & \text{if } b_k^s > 0; b_k^c < 0 \\ \pi + \tan^{-1} |b_k^s/b_k^c| & \text{if } b_k^s < 0; b_k^c < 0 \\ 2\pi - \tan^{-1} |b_k^s/b_k^c| & \text{if } b_k^s < 0; b_k^c > 0 \end{cases} \quad (3.10)$$

$$b_k = \sqrt{(b_k^s)^2 + (b_k^c)^2}$$

Now define  $\Lambda$  as the deterministic signal matrix as  $\Lambda = [\Lambda_1 | \Lambda_2]$  where the  $M$  by  $P$  matrices,  $\Lambda_1$  and  $\Lambda_2$ , are constructed as

$$\Lambda_1 = \begin{bmatrix} \cos(\omega_1[M_o - 1]) & \dots & \cos(\omega_P[M_o - 1]) \\ \vdots & \dots & \vdots \\ \cos(\omega_1[N - 1]) & \dots & \cos(\omega_P[N - 1]) \end{bmatrix} \quad (3.11)$$

while

$$\Lambda_2 = \begin{bmatrix} -\sin(\omega_1[M_o - 1]) & \dots & -\sin(\omega_P[M_o - 1]) \\ \vdots & \dots & \vdots \\ -\sin(\omega_1[N - 1]) & \dots & -\sin(\omega_P[N - 1]) \end{bmatrix} \quad (3.12)$$

With scaling vectors  $b_s = [b_1 \sin(\phi_1) \dots b_P \sin(\phi_P)]^T$  and  $b_c = [b_1 \cos(\phi_1) \dots b_P \cos(\phi_P)]^T$ , a vector form of Equation 3.9 can be written as

$$y = \Lambda_1 b_c + \Lambda_2 b_s + w = \Lambda b + w \quad (3.13)$$

where  $\mathbf{b}$  is the vector of scaling coefficients,  $\mathbf{b}^T = [\mathbf{b}_c^T; \mathbf{b}_s^T]$  and the  $2P$  columns of  $\Lambda$  are linearly independent. This is equivalent to viewing  $\mathbf{y}$  as a set of observations generated from the deterministic system  $\Lambda\mathbf{b}$  corrupted by the colored noise vector  $\mathbf{w}$ . The problem of estimating the frequencies, phases, and amplitudes has been transformed into the problem of estimating the signal matrix  $\Lambda$  and the scaling coefficients  $\mathbf{b}_c$  and  $\mathbf{b}_s$ . Once the scaling coefficients have been found, the associated phases and amplitudes can be found by Equation 3.10.

**3.2.4 Section Summary.** This section derived the mathematical model for any system required to estimate the amplitudes, phases and frequencies of filtered sinusoids. Specifically, once the steady state has been reached, the output of a filter due to an input consisting of a linear sum of  $P$  sinusoids in zero mean, independent, normally distributed noise can be represented as the output from a deterministic system,  $\Lambda\mathbf{b}$ , corrupted by zero mean, normally distributed noise with a covariance matrix,  $K_w$ . Estimation of the sinusoidal parameters has been recast as the estimation of the deterministic system  $\Lambda\mathbf{b}$ . Methods for estimating the parameters of this model will be derived in the next section.

### 3.3 Sinusoidal Parameter Estimation

This section derives algorithms for estimating the amplitudes, frequencies and phases of filtered sinusoids in noise. After reviewing estimation background in general, this section covers the principles of maximum likelihood estimation and employs these principles to obtain parameter estimators for filtered sinusoids.

**3.3.1 Estimation Background.** In general, estimating the parameters of  $P$  sinusoids in noise deals with inferring the values of the unknown parameters from a set of observations (87:279). That is, from Equation 3.9,  $M$  measurements are obtained which contain the  $P$  sinusoids embedded in noise. Since noise is a random quantity, each observation,  $y[m]$ , becomes a random variable. Consequently,  $\mathbf{y}$  becomes a random vector. By defining  $\boldsymbol{\theta}$  as a vector form of the parameters,  $\boldsymbol{\theta}^T = [\boldsymbol{\theta}_b^T, \boldsymbol{\theta}_\phi^T, \boldsymbol{\theta}_\omega^T] = [b_1 \dots b_P; \phi_1 \dots \phi_P; f_1 \dots f_P]^T$ , an estimator, or function of the data  $g_i(\cdot)$ , can be constructed which assigns a value to  $\boldsymbol{\theta}$ , as  $\hat{\boldsymbol{\theta}} = g_i(\mathbf{y})$ , for each realization of  $\mathbf{y}$ . Here,  $\hat{\boldsymbol{\theta}}$  is called the point estimate of  $\boldsymbol{\theta}$  (61). As

a function of a random vector,  $\mathbf{y}$ , the parameter vector,  $\hat{\boldsymbol{\theta}}$ , becomes a random vector with statistical characteristics defined by the Probability Density Function (PDF) of  $\mathbf{y}$  and the form of  $g_i(\cdot)$ . The accuracy of subsequent estimates are dependent primarily on the bias and variance of the estimates generated by the estimator.

Formally, an estimator  $g_i(\cdot)$  is said to produce unbiased estimates of the true parameter vector,  $\boldsymbol{\theta}$ , if the expectation of  $\hat{\boldsymbol{\theta}}$  is equal to  $\boldsymbol{\theta}$  (27:204). That is

$$E\{\hat{\boldsymbol{\theta}}\} = \boldsymbol{\theta} \quad (3.14)$$

where  $E\{\cdot\}$  denotes the expectation operator. Thus, on average, the unbiased estimator will provide the true value of  $\boldsymbol{\theta}$  (40:15-23). In addition,  $g_i(\cdot)$  is said to be the minimum variance estimator if the variance of each component of  $\hat{\boldsymbol{\theta}}$  is less than the variance produced by any other estimator (40:23). This implies  $g_i(\mathbf{y})$  is the minimum variance estimator if

$$K_{\hat{\boldsymbol{\theta}}_i} - K_{\hat{\boldsymbol{\theta}}_j} \geq [0] \quad \text{for } i \neq j \quad (3.15)$$

Here  $K_{\hat{\boldsymbol{\theta}}_i}$  denotes the covariance matrix of the estimate provided by  $g_i(\mathbf{y})$  and the term  $\geq [0]$  is interpreted as meaning the resultant matrix is positive semidefinite (40:44). If a minimum variance unbiased estimator can be found, on average, the estimates obtained will be closer to the true value than those provided by any other estimator.

A critical step in the development and analysis of point estimators is to accurately model the PDF of the data (40:7). Since  $\mathbf{y}$  is a random vector which contains the parameters to be estimated, the PDF of  $\mathbf{y}$  is said to be parameterized by  $\boldsymbol{\theta}$  for  $\boldsymbol{\theta} \in \Theta$  and written as  $p(\mathbf{y}; \boldsymbol{\theta})$  to show this dependency. Here,  $\Theta$  is called the parameter space and contains the restrictions on the allowable values of  $\boldsymbol{\theta}$  (27:201). Furthermore, since the actual value of  $\boldsymbol{\theta}$  affects the probability of observing  $\mathbf{y}$ , selection of the proper PDF describing the data is critical in deriving a good estimator. From the data model of Equation 3.13, the measurement vector,  $\mathbf{y}$ , is an  $M$ -dimensional, normally distributed random vector with a known covariance matrix resulting from the system filter matrix,  $H$ . The associated PDF of  $\mathbf{y}$  is thus parameterized by the scaling coefficients,  $\mathbf{b}$ , and the frequencies,  $\mathbf{f}$   $\mathbf{f} = [f_1 \dots f_P]^T$ .

That is

$$p(\mathbf{y}; \boldsymbol{\theta}) = [2\pi]^{-.5M} |K_{\mathbf{w}}|^{-.5} \exp\left[-\frac{1}{2}(\mathbf{y} - \Lambda \mathbf{b})^T K_{\mathbf{w}}^{-1} (\mathbf{y} - \Lambda \mathbf{b})\right] \quad (3.16)$$

where  $\boldsymbol{\theta}^T = [\mathbf{b}^T, \mathbf{f}^T]$ . The accuracy of any estimator will be strongly dependent on how well the estimator incorporates this knowledge of the PDF into its development.

**3.3.2 Estimation Accuracy.** To analyze the performance of any estimator, it is useful to have a limit or bound which indicates the best estimation accuracy any unbiased estimator may obtain from the available data (71). In particular, for sinusoidal signals embedded in zero mean, normally distributed noise, the Cramer Rao Lower Bound (CRLB) provides a limit on the accuracy any unbiased estimator,  $\hat{\boldsymbol{\theta}}$ , can attain. Hence

$$K_{\hat{\boldsymbol{\theta}}} - [\mathcal{F}(\boldsymbol{\theta})]^{-1} \geq [0] \quad (3.17)$$

Here,  $\mathcal{F}(\boldsymbol{\theta})$  is the Fisher Information matrix defined by (40:44)

$$[\mathcal{F}(\boldsymbol{\theta})]_{ij} = -E \left\{ \frac{\partial^2 \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right\} = E \left\{ \left[ \frac{\partial \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_i} \right] \left[ \frac{\partial \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_j} \right] \right\} \quad (3.18)$$

Since the observation vector,  $\mathbf{y}$ , is normally distributed with a mean,  $\mathbf{m}_{\mathbf{y}}$ , and covariance matrix,  $K_{\mathbf{w}}$ , the Fisher Information matrix can be shown to be given by (40:47)

$$[\mathcal{F}(\boldsymbol{\theta})]_{ij} = \left[ \frac{\partial \mathbf{m}_{\mathbf{y}}}{\partial \theta_i} \right]^T [K_{\mathbf{w}}]^{-1} \left[ \frac{\partial \mathbf{m}_{\mathbf{y}}}{\partial \theta_j} \right] \quad (3.19)$$

For sinusoids in noise, the CRLB will be a function of the frequencies, phases, and amplitudes present in the signal. Figure 3.1 shows the CRLB for frequency estimates, as a function of the signal frequency, for one and two sinusoids. In Figure 3.1a, for the non-filtered signal, the frequency estimation accuracy will be relatively constant except for extremely small ( $f < 0.001$ ) or extremely high ( $f > 0.499$ ) frequencies. For the filtered sinusoid, the estimation accuracy will be optimal within the filter passband. Figure 3.1b shows a contour plot of the combined CRLB for the frequencies of two sinusoids in zero-mean, normally distributed, independent noise. For this small number of measurements,

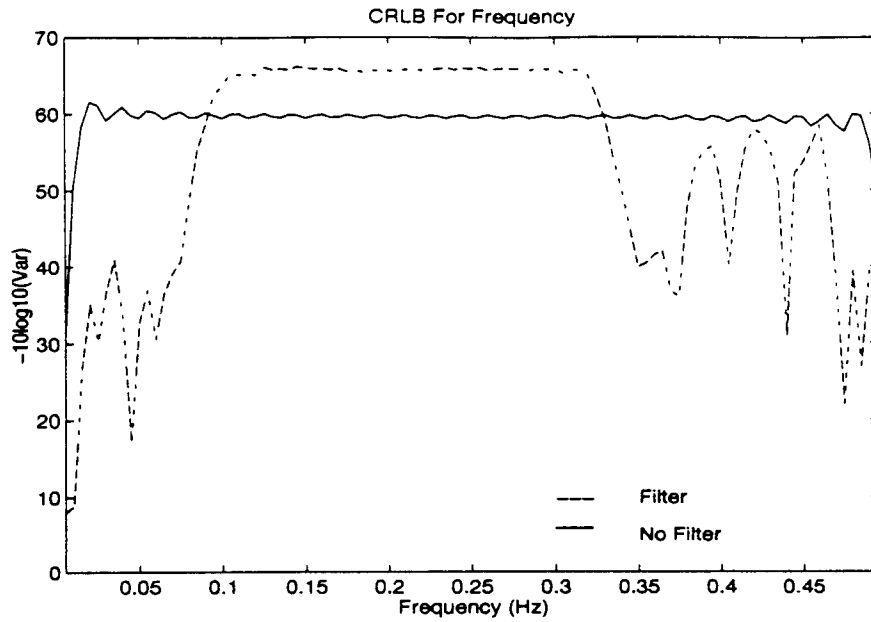
$M = 32$ , this figure indicates the estimation accuracy of any estimator will decrease dramatically for closely spaced sinusoids, ( $|f_1 - f_2| < .02$ ).

**3.3.3 Estimator Evaluation.** Since the CRLB is a complicated function of the frequencies, phases and amplitudes of the sinusoids, there is no closed form expression for determining whether an estimator which will achieve the bound for any arbitrary  $\theta$  (39:414). Consequently, the performance of any estimator must be evaluated numerically via Monte Carlo experiments (40:164-167). For this dissertation, the performance of an estimator will be determined by conducting 500 independent trials at each free variable of interest and determining the combined Mean Square Error (MSE) for each parameter to be estimated.

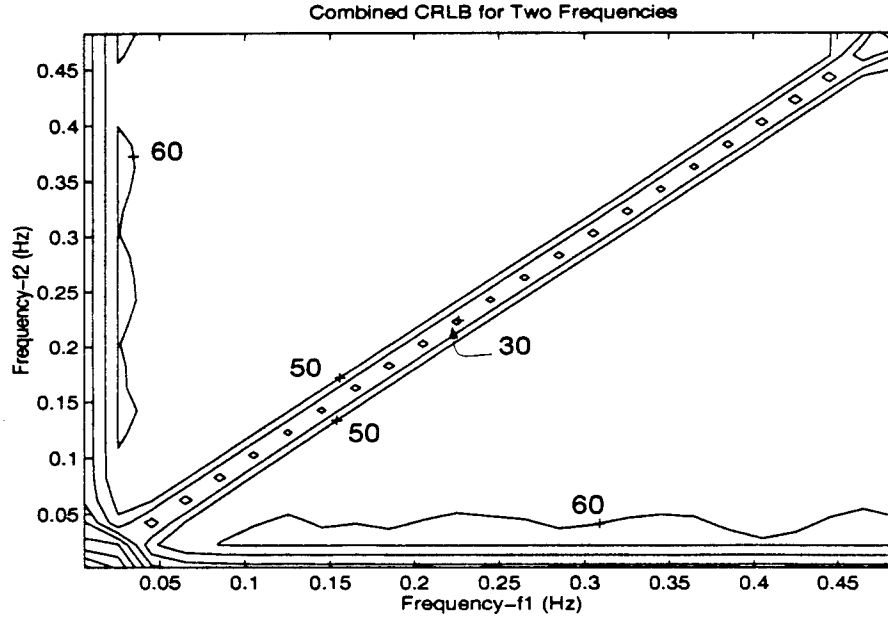
$$MSE = \frac{1}{500P} \sum_{p=1}^P \sum_{i=1}^{500} (\hat{\theta}_{p,i} - \theta_{p,i})^2 \quad (3.20)$$

The resulting experimentally obtained MSE will then be compared to the associated CRLB. Though various techniques exist to construct estimators based on the PDF given by Equation 3.16, this dissertation will employ the maximum likelihood estimation technique. The estimators provided by this technique are relatively simple to develop for signals embedded in normally distributed noise and produce the approximate minimum variance unbiased estimates for a moderate number ( $M \geq 30$ ) of observations (40:157-198).

**3.3.4 Maximum Likelihood (ML) Estimation.** The Maximum Likelihood (ML) estimation technique is based on the assumption that a particular parameter vector,  $\theta_i$ , generates different sets of observations than any other parameter vector,  $\theta_j$ . Thus, any given observation vector,  $\mathbf{y}$ , is more likely to have been generated by one particular parameter vector  $\theta_i$ , than any others (60:147). Obtaining an ML estimator involves specifying the likelihood function and finding the values of the parameter vector which maximize this function for a fixed observation vector (60:49). In general, when viewed as a function of  $\theta$  for a fixed observation vector, the PDF defined in Equation 3.16 is termed a likelihood function (40:29). Thus, for a fixed observation vector  $\mathbf{y}$ , the ML estimate for  $\theta$  maximizes  $p(\mathbf{y}; \theta)$ . Equivalently, since the natural logarithm is a monotonically increasing function, the ML estimate for  $\theta$  maximizes  $\ln\{p(\mathbf{y}; \theta)\}$ ; the log likelihood function.



(a) Sinusoid Parameters:  $[b_1^o = 1, \phi_1^o = \pi/3]$ , Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, \phi_1^o = \pi/3]$ ,  $[b_2^o = 1, \phi_2^o = 4\pi/3]$ .

Figure 3.1 **CRLB for Sinusoids in Noise:** Bound obtained for one sinusoid in (a) and two sinusoids in (b). Variance bounded as function of frequency at fixed SNR of 10dB and block length  $M = 32$ . For two sinusoids, contours plotted as  $Var = .5[Var\{f_1\} + Var\{f_2\}]$ . SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ .

Furthermore, if  $\alpha = g(\theta)$ , where  $g(*)$  is an arbitrary  $k$ -dimensional invertible function of  $\theta$ , the ML estimate for  $\alpha$  is given by

$$\hat{\alpha}_{ML} = g(\hat{\theta}_{ML}) \quad (3.21)$$

This property, known as the Invariance Property of ML estimates, states that any invertible function of an ML estimate will also produce an ML estimate (40:182).

**3.3.5 Parameter Estimators.** From the PDF of  $y$  given by Equation 3.16, with  $K_w$  assumed known, as shown in Appendix B, the ML estimate for  $\theta$  can be obtained by minimizing the following objective function,  $J(\theta)$ , with respect to the parameters in  $\theta$ :

$$J(\theta) = [y - \Lambda b]^T K_w^{-1} [y - \Lambda b] \quad (3.22)$$

Minimizing this function with respect to  $\theta$  can be accomplished by minimizing with respect to the scaling coefficients,  $b$ , and frequencies,  $f$ , individually. For a fixed ML estimate of  $f$ , the ML estimator for  $b$  is unique and found as (22), (40:186)

$$\hat{b}(ML) = [\Lambda^T K_w^{-1} \Lambda]^{-1} \Lambda^T K_w^{-1} y \quad (3.23)$$

In addition, the statistical characteristics of  $\hat{b}$  become (60)

$$E\{\hat{b}(ML)\} = b \quad \text{and} \quad K_{\hat{b}(ML)} = [\Lambda^T K_w^{-1} \Lambda]^{-1} \quad (3.24)$$

This estimator is known as the Best Linear Unbiased Estimator (BLUE) for a fixed set of frequencies (40:140). The ML estimate for  $f$ , on the other hand, minimizes the following objective function (40:186):

$$J(f) = y^T K_w^{-1} y - y^T K_w^{-1} \Lambda [\Lambda^T K_w^{-1} \Lambda]^{-1} \Lambda^T K_w^{-1} y \quad (3.25)$$

This objective function,  $J(f)$ , termed the ML frequency objective function, is a strongly nonlinear function with respect to  $f$ . Since the ML estimates of the scaling coefficients,  $b$ , are defined as a function of the frequencies,  $f$ , ML estimation of the frequencies is

critical for obtaining accurate parameter estimates. The global minimization of this non-linear function typically involves a computationally intensive search in the  $P$ -dimensional frequency domain and various techniques have been tried to accomplish this search (80).

*3.3.6 ML Frequency Estimation.* Since the term  $\mathbf{y}^T K_{\mathbf{w}}^{-1} \mathbf{y}$  is a positive number, minimizing  $J(\mathbf{f})$  can be accomplished by maximizing the following objective function

$$J_1(\mathbf{f}) = \mathbf{y}^T K_{\mathbf{w}}^{-1} \Lambda [\Lambda^T K_{\mathbf{w}}^{-1} \Lambda]^{-1} \Lambda^T K_{\mathbf{w}}^{-1} \mathbf{y} \quad (3.26)$$

One convenient method of accomplishing this maximization involves the general method of iterative gradient ascent

$$\mathbf{f}(i+1) = \mathbf{f}(i) + D^{-1} \nabla J_1(\mathbf{f}) \quad (3.27)$$

Here, each element of the gradient is found as

$$[\nabla J_1(\mathbf{f})]_p = \left[ \frac{\partial J_1(\mathbf{f})}{\partial \mathbf{f}} \right]_p = \frac{\partial J_1(\mathbf{f})}{\partial f_p} \quad (3.28)$$

and  $D$  is the Hessian of  $J_1(\mathbf{f})$ .

$$[D]_{i,j} = \left[ \frac{\partial^2 J_1(\mathbf{f})}{\partial \mathbf{f} \partial \mathbf{f}^T} \right]_{i,j} = \frac{\partial^2 J_1(\mathbf{f})}{\partial f_i \partial f_j} \quad (3.29)$$

In the area of signal processing and numerical analysis, this technique is known as the Newton-Raphson Technique (40:187) while in the area of pattern recognition, it is known as the Conjugate Gradient Method (3:274-275). Regardless of the name, as shown Appendix E, the expression for the gradient becomes relatively simple while the calculation of the Hessian is computationally intensive. An alternative method of maximizing  $J_1(\mathbf{f})$  can be found by using the direct gradient ascent maximization technique (7)

$$\mathbf{f}(i+1) = \mathbf{f}(i) + \eta(i) \nabla J(\mathbf{f}) + \gamma(i) [\mathbf{f}(i) - \mathbf{f}(i-1)] \quad (3.30)$$

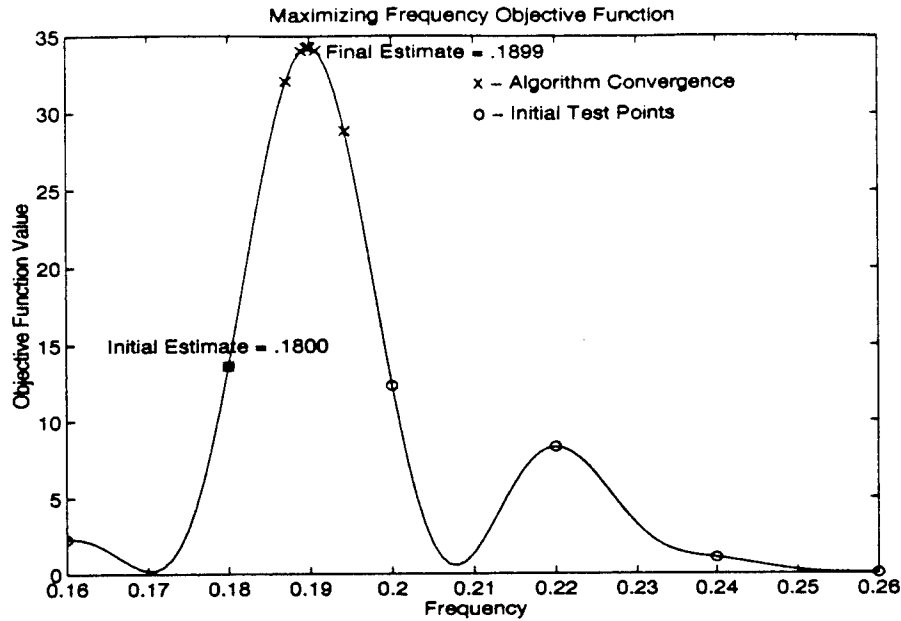


Here, the rate of convergence is controlled by the adaptive gain term  $\eta(i)$  and the adaptive momentum term  $\gamma(i)$ . As with most iterative optimization techniques, there is no guarantee the algorithm will ever converge. Even if it does converge, there is no guarantee the global maximum has been attained, especially for nonconvex functions. The solution achieved will only provide a local maximum and the accuracy of the estimate will depend greatly on the location of the initial guess.

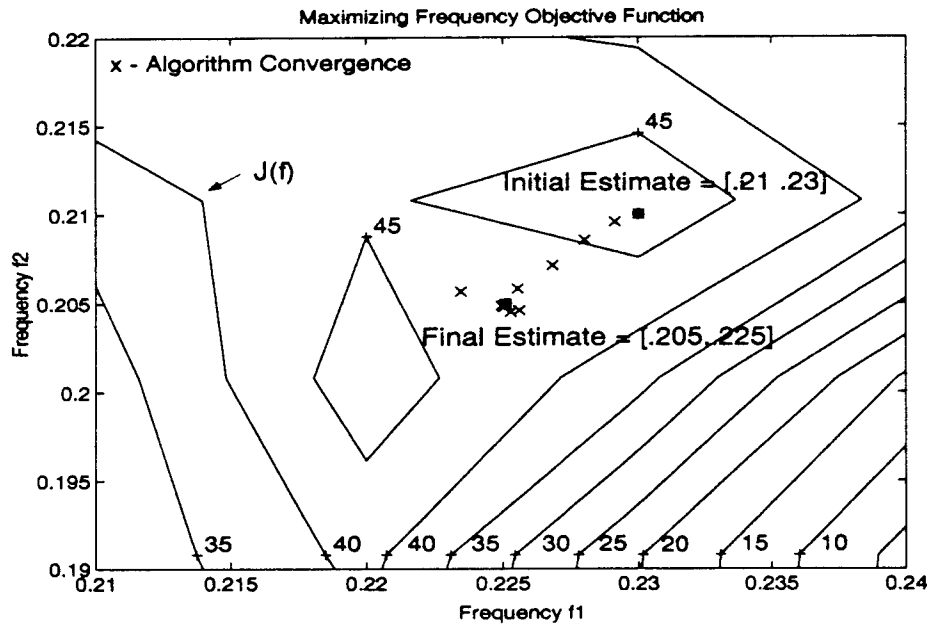
Figure 3.2 shows the convergence of this iterative technique for one and two filtered sinusoids. Since the objective function,  $J_1(\mathbf{f})$ , contains more than one local maximum, the initial estimate was obtained by evaluating  $J_1(\mathbf{f})$  at equally spaced test points and choosing the frequency with the largest  $J_1(\mathbf{f})$ . In both instances, the method of direct gradient ascent found the global maximum in less than 30 iterations.

Figure 3.3, Figure 3.4, and Figure 3.5 show the estimation accuracy attained using ML estimates of the parameters of one and two filtered sinusoids. This figure shows the ML parameter estimates will achieve the CRLB, above threshold, only when the correlation in the noise is correctly modeled. Assumptions that the noise is uncorrelated (white) leads to an incorrect model and suboptimal estimates. Since most digital signal processing systems incorporate an antialiasing filter as a minimum, the noise will always be correlated. The results of these figures show the fallacy of using a white noise model to construct the CRLB for filtered signals.

*3.3.7 Section Summary.* This section derived ML estimators for the amplitude, phase and frequency parameters of  $P$  filtered sinusoids in noise and represents an original contribution of this research. Specifically, by correctly accounting for the correlation in the noise due to the effects of the filter, an accurate PDF model of the measurements was derived. Manipulation of the objective function derived from this PDF showed ML frequency estimates provide ML estimates of the sinusoidal amplitudes and phases. Simulations indicate the method of gradient ascent can be used to provide the ML frequency estimates provided the PDF model accurately describes the correlation in the noise due to the filter. Failure to incorporate the filter effects in the ML model will lead to subopti-

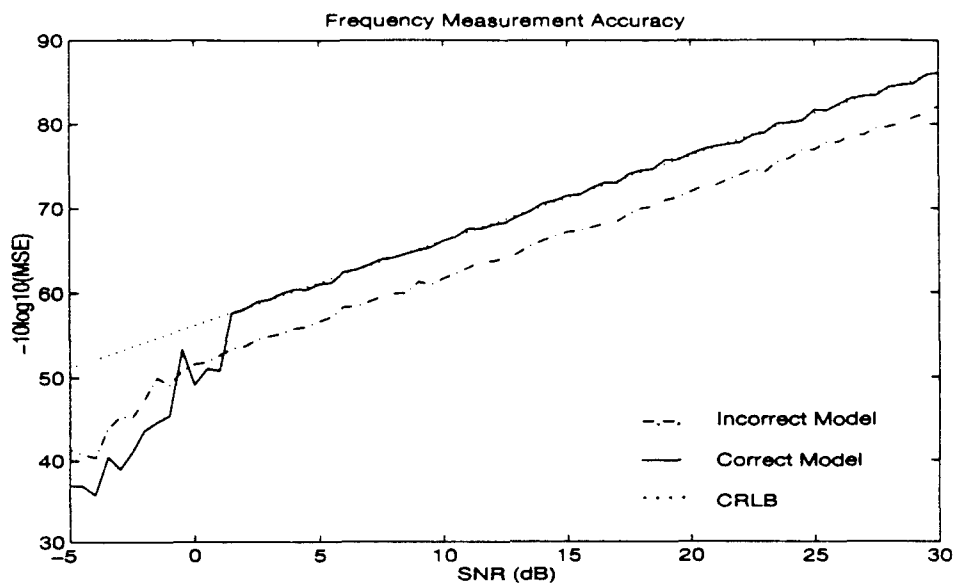


(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .191, \phi_1^o = 3\pi/8]$ ,  $SNR = 0\text{dB}$ .

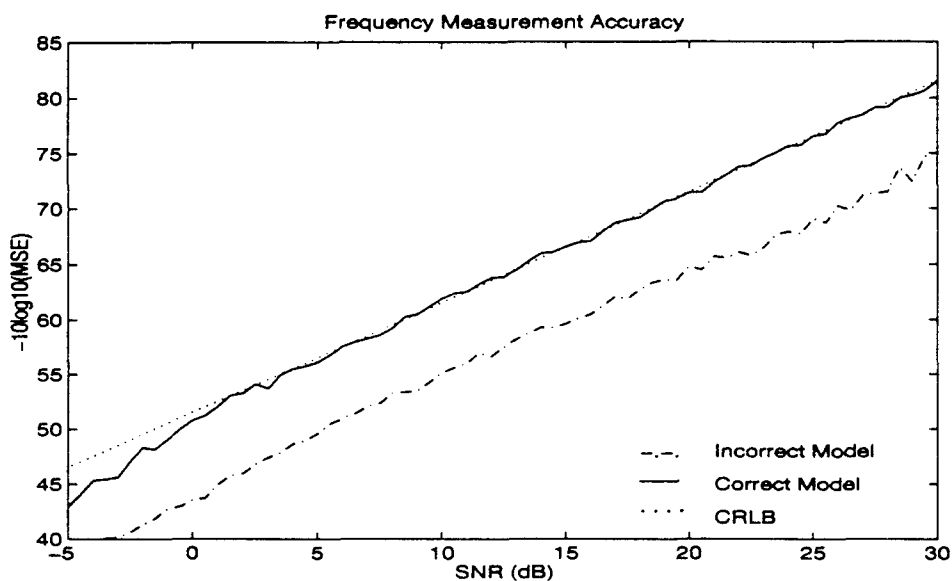


(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = 3\pi/8]$ ,  $[b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$  and  $SNR = .5\text{dB}$ .

Figure 3.2 **Maximizing ML Frequency Objective Function:**  $J(f)$  calculated as function of frequency at fixed SNR and block length  $M = 32$ . Initial estimates made by choosing  $f$  with largest  $J(f)$  from uniform samples across filter bandwidth. Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ ; SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ .

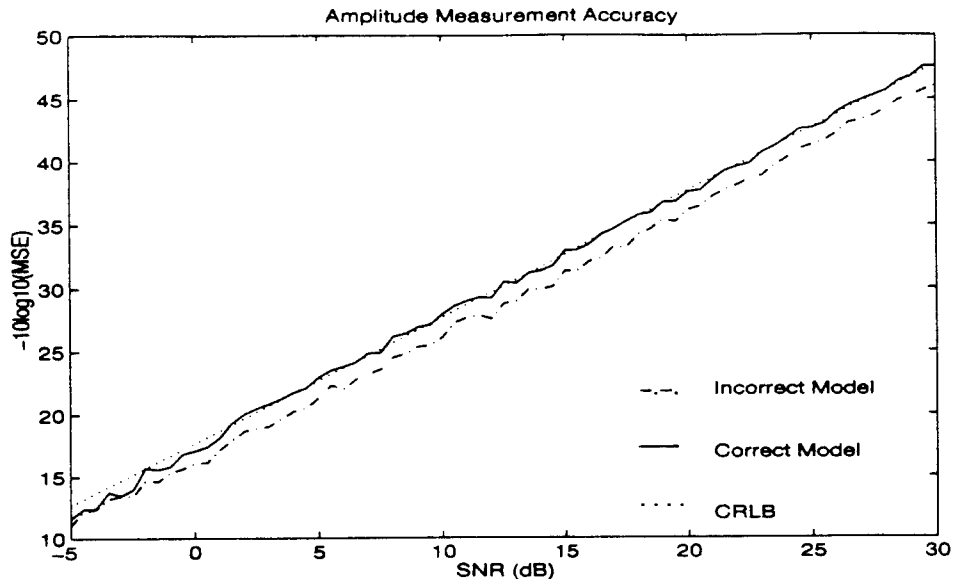


(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ .

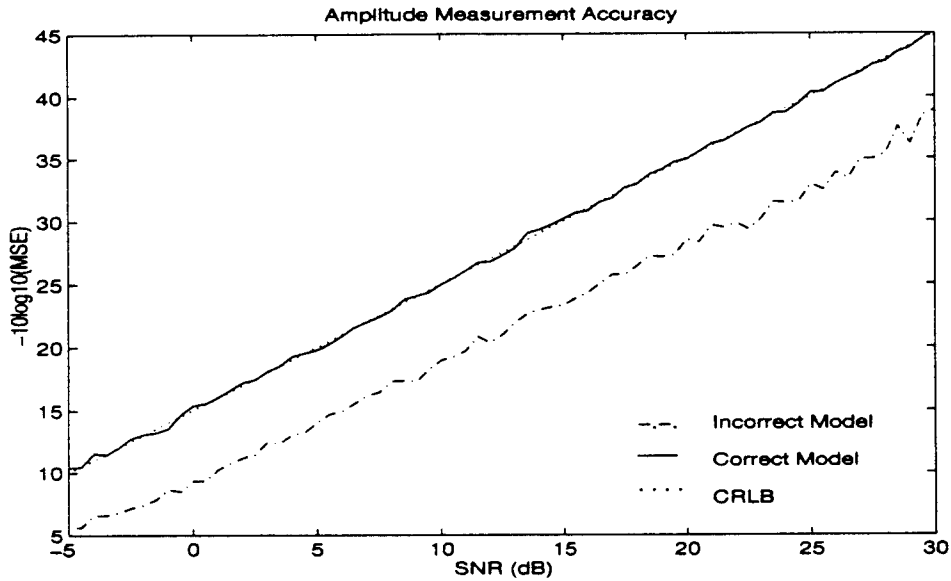


(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3], [b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .

**Figure 3.3 ML Frequency Estimation Accuracy:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. The incorrect model assumed  $K_w = \sigma^2 I$  while the correct model used  $K_w = \sigma^2 H^T H$ ; Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ ; SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ .

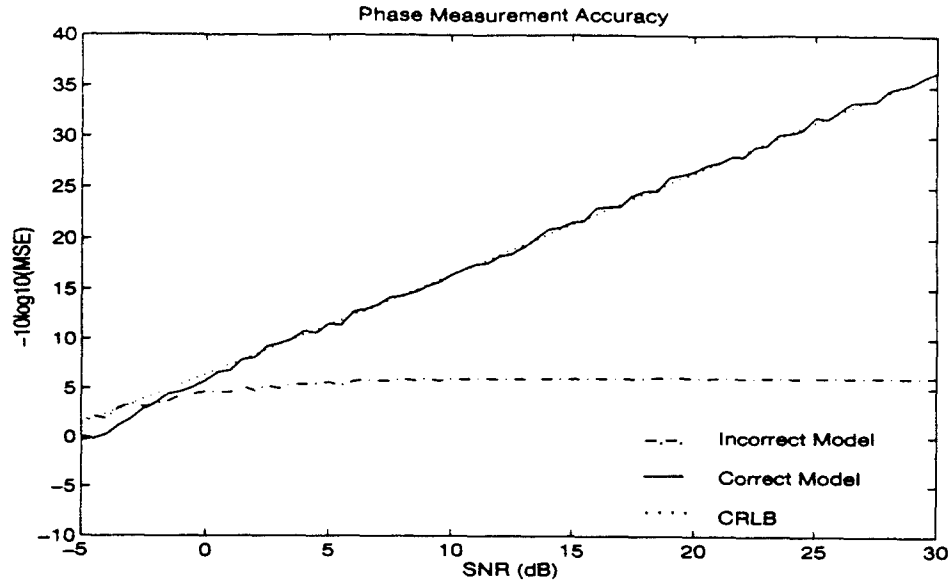


(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ .

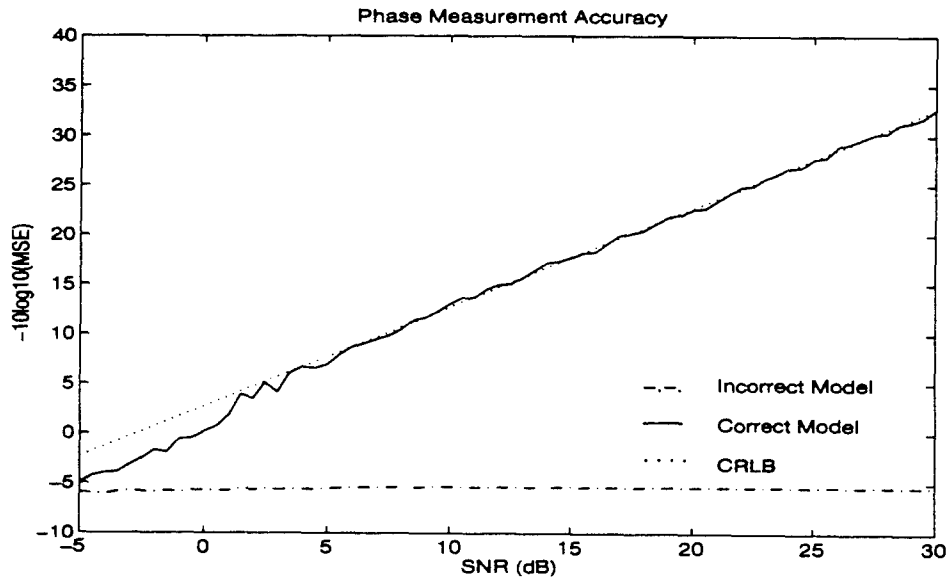


(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3], [b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .

**Figure 3.4 ML Amplitude Estimation Accuracy:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. The incorrect model assumed  $K_w = \sigma^2 I$  while the correct model used  $K_w = \sigma^2 H^T H$ ; Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ ; SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$  and  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ .



(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3], [b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .

**Figure 3.5 ML Phase Estimation Accuracy:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. The incorrect model assumed  $K_w = \sigma^2 I$  while the correct model used  $K_w = \sigma^2 H^T H$ ; Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ ; SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$  and  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ .

mal parameter estimators. The next section introduces an alternate method for frequency estimation based on linear prediction modeling.

### 3.4 Linear Prediction (LP) Modeling

This section develops an alternate method of estimating the frequencies of pulsed sinusoids based on Linear Prediction (LP) modeling techniques. Specifically, this section derives the true linear model, parameterized by the LP coefficients, for estimating the frequencies of  $P$  sinusoids in noise and is an original contribution of this research. Using this model and properly accounting for the effects of the LP coefficients and the system filter on the noise, estimation of the frequencies is recast as the estimation of the LP coefficients. This section concludes by developing a bound on the LP coefficient estimation error.

*3.4.1 Theoretical Background.* To show the relationship between sinusoidal frequencies and a set of LP coefficients, consider a signal  $s[m]$ , consisting of the sum of  $P$  real sinusoids. As shown in Chapter II, there exists a set of  $2P + 1$  LP coefficients,  $a[p]$ , such that (8)

$$\sum_{p=0}^{2P} a[p]s[m-p] = 0$$

Since each frequency of  $s[m]$  is a zero of the polynomial,  $A(z)$ , formed from the LP coefficients, the  $2P$  frequencies present in  $s[m]$  can be found as roots to  $A(z)$ . The search for the  $P$  frequencies comprising the sinusoids in  $s[m]$  has been recast as the search for the associated  $2P + 1$  LP coefficients. As shown in Chapter II, many algorithms based on estimating these coefficients have been developed in the past. This dissertation employs stochastic modeling techniques to derive a new method of LP modeling which is based on the general linear system model (40:94-94)

$$t = U\phi + v \tag{3.31}$$

In this model,  $t$  is an  $M$ -dimensional vector called the measurement vector,  $\phi$  is a  $P$ -dimensional vector called the parameter vector,  $U$  is an  $M$  by  $P$  matrix called the observation matrix and  $v$  is called the measurement error vector (60:10).

**3.4.2 Construction of the LP General Linear Model.** A convenient method of calculating the LP coefficients can be found by modeling Equation 2.14 in terms of the general linear system model of Equation 3.31. Since Equation 2.14 must hold for  $2P \leq m \leq M - 1$ , a deterministic linear model, parameterized by the LP coefficients in the vector  $\mathbf{a} = [a[0] \dots a[2P]]^T$ , can be constructed as

$$S\mathbf{a} = \begin{bmatrix} s[M-1] & \dots & s[M-1-2P] \\ \vdots & \dots & \vdots \\ s[2P] & \dots & s[0] \end{bmatrix} \begin{bmatrix} a[0] \\ \vdots \\ a[2P] \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.32)$$

or equivalently  $S\mathbf{a} = \mathbf{0}$ . Here, the  $M - 2P$  by  $2P + 1$  matrix  $S$  is deterministic with a rank of  $2P$  so that  $\mathbf{a}$  resides in the null space of  $S$  (45). Imposing the constraint,  $a[0] = 1$  yields the deterministic general linear model

$$\mathbf{s}_o = -S_o \mathbf{a}_o \quad (3.33)$$

Here,  $\mathbf{a}_o = [a[1] \dots a[2P]]^T$ , while  $\mathbf{s}_o$  is the first column of  $S$  and  $S_o$  is the observation matrix of the remaining columns of  $S$ . The vector,  $\mathbf{s}_o$ , can now be thought of as being generated from the deterministic matrix,  $-S_o$ , so that only  $4P + 1$  samples are needed to determine  $\mathbf{a}_o$  exactly. Unfortunately,  $s[m]$  is not usually available and estimates of  $\mathbf{a}_o$  must be obtained using the noisy measurements  $y[m]$ .

**3.4.3 Noise Effects in the Measurements.** To study the affects of noise in the measurements, assume  $S_o$  is known exactly while the elements in  $\mathbf{s}_o$  are corrupted with additive noise,  $w[m]$ , so that  $s[m] = y[m] - w[m]$  for  $m = 2P \dots M - 1$ . In this case, the form of the LP general linear model becomes

$$\mathbf{y}_o = -S_o \mathbf{a}_o + \mathbf{w}_o = \mathbf{s}_o + \mathbf{w}_o \quad (3.34)$$

Here,  $w_o$  can be viewed as the error between the measurement vector,  $y_o$ , and the vector  $s_o$  generated by the LP general linear model. The least squares estimate for  $a_o$  which minimizes the sum of the squared errors is given as (60:31)

$$\hat{a}(LS) = - [S_o^T S_o]^{-1} S_o^T y_o \quad (3.35)$$

A basic assumption for this solution is that the impact of the error in each measurement,  $y[m]$ , is the same (83:203). When this is not the case, a weight matrix,  $C^{-1}$ , can be inserted to provide the Weighted Least Squares (WLS) estimate (60:31),(40:141).

$$\hat{a}(WLS) = - [S_o^T C^{-1} S_o]^{-1} S_o^T C^{-1} y_o \quad (3.36)$$

For the data model given by Equation 3.13, the noise is assumed to be zero mean and normally distributed with covariance matrix  $K_w$ . Thus, the measurement vector  $y_o$  is normally distributed with the following properties

$$\text{Mean} = m_{y_o} = -S_o a_o \quad (3.37)$$

$$\text{Covariance} = K_{y_o} = K_w \quad (3.38)$$

As such, the ML estimate for the LP coefficients is given by (60:152-155),(40:185-187)

$$a_o(ML) = - [S_o^T K_w^{-1} S_o]^{-1} S_o^T K_w^{-1} y_o \quad (3.39)$$

This estimate is unbiased with a covariance matrix given by (60:110-113)

$$K_{a_o(ML)} = [S_o^T K_w^{-1} S_o]^{-1} \quad (3.40)$$

If the deterministic observation matrix,  $S_o$ , was known, the ML estimate obtained in Equation 3.39 would represent the minimum variance, unbiased estimate obtainable from the measurements. Unfortunately, the observation matrix must also be estimated from the measurements.



3.4.4 *Noise Effects in the Observation Matrix.* The main drawback in using the LP general linear model to identify the LP coefficients is that the effects of the noise are embedded in the observation matrix. That is, since  $s[m] = y[m] - w[m]$ , the observation matrix itself,  $S_o$ , is perturbed by the noise. In terms of the true LP coefficients, the LP general linear model becomes

$$\mathbf{y}_o = -Y_o \mathbf{a}_o + \mathbf{w}_o + W_o \mathbf{a}_o \quad (3.41)$$

where the matrices  $Y_o$  and  $W_o$  are defined in a manner similar to  $S_o$ . Defining the error as

$$\mathbf{z}(\mathbf{a}) = \mathbf{w}_o + W_o \mathbf{a}_o = W \mathbf{a} \quad (3.42)$$

the perturbed form of the LP general linear model becomes

$$\mathbf{y}_o = -Y_o \mathbf{a}_o + \mathbf{z}(\mathbf{a}) \quad (3.43)$$

The error vector,  $\mathbf{z}(\mathbf{a})$ , now has statistics which are dependent on the LP coefficients in addition to the statistics of the noise vector  $\mathbf{w}$ . To show this dependency, define  $A$  as the  $M$  by  $M - 2P$  matrix of LP coefficients so that

$$A^T = \begin{bmatrix} a[0] & \dots & a[2P] & 0 & 0 & \dots & 0 \\ 0 & a[0] & \dots & a[2P] & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & \vdots & 0 \\ 0 & 0 & \dots & 0 & a[0] & \dots & a[2P] \end{bmatrix} \quad (3.44)$$

This matrix has a rank of  $M - 2P$  and allows the error vector,  $\mathbf{z}(\mathbf{a})$  to be written as

$$\mathbf{z}(\mathbf{a}) = \mathbf{w}_o + W_o \mathbf{a}_o = W \mathbf{a} = A^T \mathbf{w} \quad (3.45)$$

Since  $\mathbf{w}$  is a zero-mean, normally distributed noise vector with covariance  $K_w$ , the error vector is a zero-mean, normally distributed random vector with a covariance matrix,  $K_{\mathbf{z}(\mathbf{a})}$ , given by

$$K_{\mathbf{z}(\mathbf{a})} = A^T K_w A = \sigma^2 A^T H^T H A \quad (3.46)$$

In order to account for this colorization in the covariance matrix, let  $G(\mathbf{a})$  be a square  $M - 2P$  matrix derived from the Cholesky decomposition of the error covariance matrix (73:440-441).

$$G^T(\mathbf{a})G(\mathbf{a}) = \frac{1}{\sigma^2} [A^T H^T H A]^{-1} = K_{\mathbf{z}(\mathbf{a})}^{-1} \quad (3.47)$$

Premultiplying Equation 3.43 by  $G(\mathbf{a})$  yields a new form of the LP general linear model

$$G(\mathbf{a})\mathbf{y}_o = -G(\mathbf{a})Y_o\mathbf{a}_o + G(\mathbf{a})\mathbf{z}(\mathbf{a}) = -G(\mathbf{a})Y_o\mathbf{a}_o + \mathbf{e}(\mathbf{a}) \quad (3.48)$$

Here,  $\mathbf{e}(\mathbf{a})$  is a zero mean, normally distributed random vector with covariance matrix  $K_{\mathbf{e}(\mathbf{a})} = I$ . The form of the LP general linear model now becomes

$$G(\mathbf{a})Y\mathbf{a} = \mathbf{e}(\mathbf{a}) \quad (3.49)$$

The estimation of the  $P$  frequencies in  $s[m]$  from Equation 3.13 has been recast as the estimation of the  $2P + 1$  LP coefficients using the LP general linear model derived in Equation 3.49. The number of LP coefficients to be estimated can be reduced further by imposing sinusoidal constraints on the coefficients.

**3.4.5 Imposition of Sinusoidal Constraints.** To properly estimate  $P$  real sinusoids using the  $2P^{\text{th}}$  order polynomial,  $A(z)$ , as shown in Appendix C, the LP coefficients must be constrained to be symmetric about  $P$  so that  $a[p] = a[2P - p]$  for  $p = 0 \dots P$  (8). To impose this constraint, let  $\boldsymbol{\alpha} = [a[0], \dots, a[P]]^T$  be a vector containing the constrained LP coefficients. Thus,  $\mathbf{a}$  is related to  $\boldsymbol{\alpha}$  via the linear transformation  $\mathbf{a} = B\boldsymbol{\alpha}$ . Here  $B$  is the constraint matrix defined as

$$B = \begin{bmatrix} I_P & \mathbf{o}_P \\ \mathbf{o}_P^T & 1 \\ IB_P & \mathbf{o}_P \end{bmatrix} \quad (3.50)$$

and  $\mathbf{0}_P$  represents a vector of  $P$  zeros,  $I_P$  represents the  $P$  square identity matrix, while  $IB_P$  represents the  $P$  square 'Backward' identity matrix given by

$$[IB_P]_{i,j} = \begin{cases} 1 & \text{for } j = P + 1 - i \\ 0 & \text{otherwise} \end{cases} \quad (3.51)$$

The search for the LP coefficients, contained in  $\mathbf{a}$ , of a  $2P^{th}$  order model has been reduced to the identification of the  $P + 1$  constrained coefficients contained in  $\boldsymbol{\alpha}$ . Imposing these constraints on the LP general linear model yields

$$G(\mathbf{a})YB\boldsymbol{\alpha} = G(\boldsymbol{\alpha})Y_C\boldsymbol{\alpha} = \mathbf{e}(\boldsymbol{\alpha}) \quad (3.52)$$

Here,  $Y_C$  is an  $M - 2P$  by  $P + 1$  matrix of rank  $P + 1$ . Now let  $\mathbf{y}_c$  be the first column of  $Y_C$  and let  $Y_c$  be an  $M - 2P$  by  $P$  matrix of the remaining columns of  $Y_C$ . Defining  $\mathbf{y}_c(\boldsymbol{\alpha}) = G(\boldsymbol{\alpha})\mathbf{y}_c$  and  $Y_c(\boldsymbol{\alpha}) = G(\boldsymbol{\alpha})Y_c$  while letting  $\boldsymbol{\alpha}_o = [a[1] \dots a[P]]^T$ , an alternate form of the LP general linear model becomes

$$\mathbf{y}_c(\boldsymbol{\alpha}) = -Y_c(\boldsymbol{\alpha})\boldsymbol{\alpha}_o + \mathbf{e}(\boldsymbol{\alpha}) \quad (3.53)$$

Equation 3.53 represents the true general linear model relating the LP coefficients to the frequencies present in  $P$  sinusoids in noise. The problem now is to estimate the LP parameters in  $\boldsymbol{\alpha}_o$  and, consequently  $G(\boldsymbol{\alpha})$ , given the  $M$  observations. Before developing any estimator for the LP coefficients, it is useful to have a bound on the accuracy that can be obtained from the measurements.

**3.4.6 Bounding the Estimation Error.** In general, let  $\boldsymbol{\alpha}$  be a function of a random vector  $\boldsymbol{\theta}$  so that  $\boldsymbol{\alpha} = g(\boldsymbol{\theta})$ . It can be shown the CRLB for  $\boldsymbol{\alpha}$  can be derived from the CRLB of  $\boldsymbol{\theta}$  as (40:45)

$$K_{\boldsymbol{\alpha}} - \left[ \frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] [\mathcal{F}(\boldsymbol{\theta})]^{-1} \left[ \frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^T \geq [0] \quad (3.54)$$

Again,  $\mathcal{F}(\theta)$  is the Fisher Information matrix for  $\theta$  while

$$\left[ \frac{\partial g(\theta)}{\partial \theta} \right]_{i,j} = \frac{\partial g_i(\theta)}{\partial \theta_j} \quad (3.55)$$

For the LP coefficients, the functions relating the  $P$  unique coefficients to be estimated to the corresponding frequencies can be found by expanding the LP polynomial

$$\sum_{p=0}^P a[p]z^{-p} = \prod_{p=1}^P [z^{-2} - 2 \cos(2\pi f_p)z^{-1} + 1] \quad (3.56)$$

In particular, for  $P = 1$

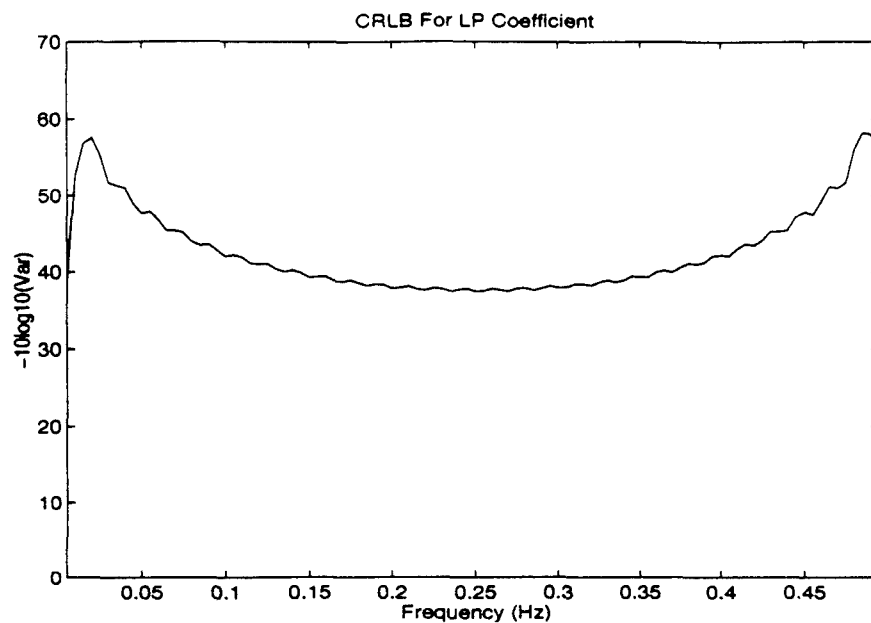
$$a[1] = -2 \cos(2\pi f_1) = g_1(f_1) \quad (3.57)$$

while for  $P = 2$

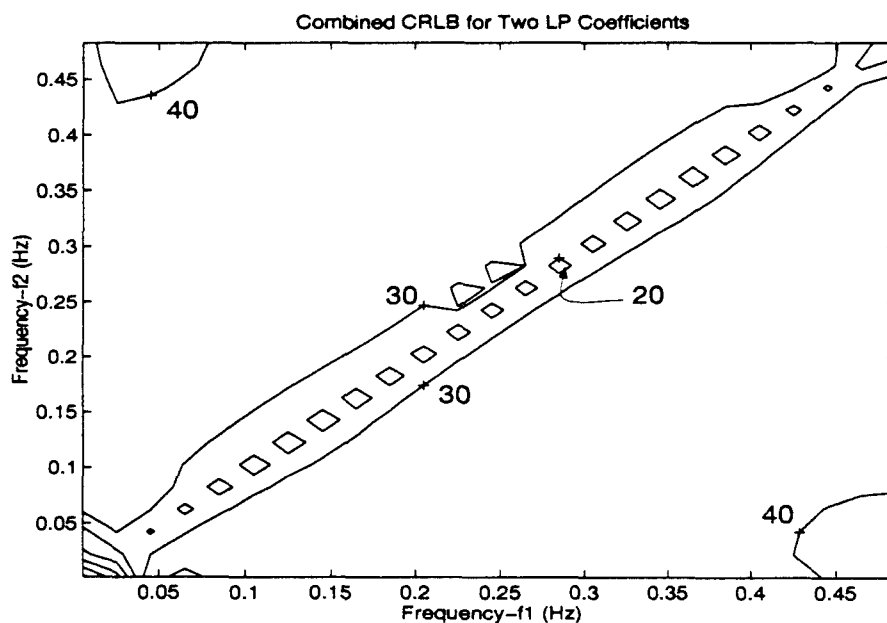
$$\begin{aligned} a[1] &= -2 \cos(2\pi f_1) - 2 \cos(2\pi f_2) = g_1(\mathbf{f}) \\ a[2] &= 2 + 4 \cos(2\pi f_1) \cos(2\pi f_2) = g_2(\mathbf{f}) \end{aligned} \quad (3.58)$$

Figure 3.6 shows the CRLB for the LP coefficients for one and two sinusoids. For one sinusoid, the estimation accuracy will decrease as the sinusoidal frequency approaches .25. This is to be expected since small changes in the frequency around  $f = .25$  will produce large changes in the value of the LP coefficient. For two sinusoids, as shown by the contour plot of Figure 3.6b, the estimation accuracy of the LP coefficients is primarily a function of the frequency separation of the sinusoids. Consequently, the estimation accuracy decreases dramatically for  $|f_1 - f_2| < .02$ . This is to be expected since the CRLB for the frequencies follows the same general form as the CRLB of the coefficients.

**3.4.7 Section Summary.** This section derived the true LP general linear model for estimating the set of  $P$  unique LP coefficients generating  $P$  sinusoids in noise and is an original contribution of this research. Specifically, by properly incorporating the effects of both the filter and the LP coefficients upon the noise, the LP general linear model derived in Equation 3.53 represents the true linear model for estimating the LP coefficients. Any other model relating the estimation of the LP coefficients to the sinusoidal frequencies will



(a) Sinusoid Parameters:  $[b_1^o = 1, \phi_1^o = \pi/3]$ , Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$  Length:  $N = 32$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, \phi_1^o = \pi/3]$ ,  $[b_2^o = 1, \phi_2^o = 4\pi/3]$ .

Figure 3.6 **CRLB for LP Coefficients:** Variance bounded as function of frequency at fixed SNR of 10dB and block length  $M = 32$ . For two LP coefficients, contours found as  $Var = .5(Var\{a[1]\} + Var\{a[2]\})$  and plotted as  $-10\log_{10}(Var)$ . SNR calculated as  $SNR = -10\log_{10}(2\sigma^2)$ .

provide suboptimal frequency estimates. Estimating the frequencies has been recast as estimating the coefficients of the LP general linear model. This derivation naturally leads to the employment of a Fixed Point concept to determine the LP coefficients. This is one of the main original contributions of this dissertation.

### 3.5 LP Coefficient Estimation: Fixed Point Concept

This section derives the estimators to be used for estimating the LP coefficients as defined by the LP general linear model derived in Equation 3.53. By viewing the optimal estimates as fixed points in the domain of allowable solutions, this section develops two alternate methods of estimating the LP coefficients. One estimator is based on an Iterative Least Squares (ILS) technique and the other on an Iterative Total Least Squares (ITLS) technique. Both of these estimators provide accurate estimates of the LP coefficients over a wide range of SNRs and represent original contributions in the area of LP parameter estimation.

**3.5.1 Theoretical Background.** The technique to be used for estimating the LP coefficients from the general linear model of Equation 3.53 is based on the theory of fixed points and reviewed in Appendix D. To apply this theory to the estimation process, a nonempty set,  $\mathcal{S}$ , must be defined along with a function,  $d$ , which computes the distance between any two members of  $\mathcal{S}$  (33:1). Assuming the properties of a general distance function hold, then the pair  $[\mathcal{S}, d]$  is called a metric space (58:21). Now let  $\mathcal{L} : \mathcal{S} \rightarrow \mathcal{S}$  be a function from a metric space  $[\mathcal{S}, d]$  into itself. A point  $\alpha \in \mathcal{S}$  is called a fixed point of  $\mathcal{L}$  provided (2:92)

$$\mathcal{L}(\alpha) = \alpha \tag{3.59}$$

In general, there are many requirements which must be satisfied for a fixed point of a function to exist (33). However, the following two theorems will guarantee the existence of a fixed point of an arbitrary nonlinear mapping,  $\mathcal{L}$ , provided the metric space is adequately defined:

- Schauder's Fixed Point Theorem - Let  $\mathcal{L}$  be a mapping on a Banach space,  $\mathcal{S}$  such that  $\mathcal{L} : \mathcal{S} \rightarrow \mathcal{S}$  is continuous and bounded and the image of each bounded set has a compact closure. Then there exists in  $\mathcal{S}$  a fixed point  $\alpha$  of  $\mathcal{L}$  (33:152).
- Brouwer's Fixed Point Theorem- Let  $\mathcal{L}$  be a continuous mapping on  $S_L(0, 1)$ , the unit sphere in  $L$ -dimensional Euclidean space. If  $\mathcal{L}(\alpha_i) \in S_L(0, 1)$  for all  $\alpha_i \in S_L(0, 1)$ , there exists a fixed point of  $\mathcal{L}$ , denoted  $\alpha$  in  $S_L(0, 1)$  (33:116).

In this dissertation, the metric space to be employed for estimating the LP coefficients will be subsets of the general  $L$ -dimensional Euclidean space.

**3.5.2 Metric Space Definition.** Define  $S_L(0, k)$  as the set of  $L$ -dimensional vectors,  $\alpha = [\alpha_1 \dots \alpha_L]^T$ , with  $\alpha_i \in R$ , such that the two-norm of  $\alpha$  is finite.

$$\|\alpha\|^2 = \alpha^T \alpha = \sum_{i=1}^L \alpha_i^2 = k < \infty \quad (3.60)$$

This implies each component of  $\alpha$  is finite. Thus,  $S_L(0, k)$  is a closed and bounded set of  $R^L$ , the Euclidean  $L$  space (58). Now define a distance function,  $d(\alpha_i, \alpha_j)$  for any  $\alpha_i, \alpha_j \in S_L(0, k)$  as

$$d(\alpha_i, \alpha_j) = \sqrt{\sum_{i=1}^L (\alpha_{ii} - \alpha_{ji})^2} \quad (3.61)$$

The pair,  $[S_L(0, k), d]$ , is a metric space. In addition, under this distance metric,  $R^L$  is a Banach space (58:112) and  $S_L(0, k)$  is called compact (58:62). Provided a function  $\mathcal{L}$  can be constructed which maps  $\alpha_i \in S_L(0, k)$  into  $\alpha_j \in S_L(0, k)$ , then a fixed point of  $\mathcal{L}$  will exist. The functions to be constructed will be based on the LS and TLS solutions to the over determined system of equations given by

$$Y_c(\alpha)\alpha_o = -y_c(\alpha) \quad (3.62)$$

### 3.5.3 Iterative Least Squares Fixed Point Mapping.

**3.5.3.1 Theory.** In general, the LS solution to Equation 3.62 consists of finding the set of coefficients,  $\alpha_o$ , such that the sum of the squared terms of the error

vector,  $e(\alpha_o)$ , defined by

$$e(\alpha_o) = Y_c(\alpha_o)\alpha_o + y_c(\alpha_o) \quad (3.63)$$

is minimized (83:154). Thus, the optimal solution is given as

$$\hat{\alpha}_o = \arg(\min_{\alpha_o} \{e(\alpha_o)^T e(\alpha_o)\}) \quad (3.64)$$

This is the same as locating the point  $\hat{y}_c(\alpha_o)$  in the column space of  $Y_c(\alpha_o)$  which is closer to  $y_c(\alpha_o)$  than any other point (83:154). Thus,  $e(\alpha_o)$  must be perpendicular to the columns of  $Y_c(\alpha_o)$  so that  $Y_c^T(\alpha_o)e(\alpha_o) = 0$  (87:524). Substitution yields the least squares estimate derived from the observations as

$$\hat{\alpha}_o = -[Y_c^T(\alpha_o)Y_c(\alpha_o)]^{-1} Y_c^T(\alpha_o)y_c(\alpha_o) = \mathcal{L}_{LS}(\alpha_o) \quad (3.65)$$

This is equivalent to finding the error vector,  $e(\alpha_o)$ , of minimum norm which must be added to the vector,  $-y_c(\alpha_o)$ , to bring it into the range of  $Y_c(\alpha_o)$  (30), (67). To use this function as an estimator of the LP coefficients, a fixed point must exist.

**3.5.3.2 Existence.** To show a fixed point exists for the LS mapping function,  $\mathcal{L}_{LS}$ , let  $\alpha_l \in S_P(0, k)$ . The LS mapping of  $\alpha_l$  yields

$$\alpha_j = \mathcal{L}_{LS}(\alpha_l) = -[Y_c^T G^T(\alpha_l)G(\alpha_l)Y_c]^{-1} [Y_c^T G^T(\alpha_l)G(\alpha_l)y_c] \quad (3.66)$$

To show  $\alpha_j \in S_P(0, k)$ , rewrite  $G(\alpha_l)$  in terms of its SVD as

$$G(\alpha_l) = U\Sigma V^T \quad (3.67)$$

where  $U = [u_1 \dots u_{M-2P}]$ ,  $V = [v_1 \dots v_{M-2P}]$  are square unitary matrices and  $\Sigma$  is an  $M - 2P$  square diagonal matrix containing the singular values of  $G(\alpha_l)$  arranged so that  $\sigma_i > \sigma_j$  for  $i > j$ . Now,  $\sigma_i = \lambda_i^{-1/2}$  where  $\lambda_i$  is an eigenvalue of the nonsingular matrix



$A_i^T K_w A_i$ . Since,  $\lambda_i > 0$  for  $i = 1 \dots M - 2P$ , the vector,  $\mathbf{x} = G(\alpha_l) \mathbf{y}_c$ , can be written as

$$\mathbf{x} = \sum_{m=1}^{M-2P} \lambda_m^{-1/2} \mathbf{u}_m \mathbf{v}_m^T \mathbf{y}_c \quad (3.68)$$

Here,  $\mathbf{x}^T \mathbf{x} < \infty$  so that  $\mathbf{x} \in \mathcal{S}_P(0, k)$ . Now, assuming  $Y_c$  is of rank  $P$ , then the matrix product,  $Y_c(\alpha_l) = Y_c G(\alpha_l)$  is also of rank  $P$ . Decomposing via the SVD yields

$$Y_c(\alpha_l) = EDQ^T \quad (3.69)$$

where  $E = [\mathbf{e}_1 \dots \mathbf{e}_{M-2P}]$  and  $Q = [\mathbf{q}_1 \dots \mathbf{q}_P]$  are square unitary matrices and  $D$  is an  $M - 2P$  by  $P$  diagonal matrix containing the nonzero singular values,  $d_p$ , for  $p = 1 \dots P$ , of  $Y_c(\alpha_l)$ . Thus

$$\|\mathcal{L}_{LS}(\alpha_l)\|^2 = \alpha_j^T \alpha_j = \sum_{p=1}^P d_p^{-2} \mathbf{x}^T \mathbf{e}_p \mathbf{e}_p^T \mathbf{x} \leq \sum_{p=1}^P d_p^{-2} \|\mathbf{x}\|^2 < \infty \quad (3.70)$$

and for any  $\alpha_l \in \mathcal{S}_P(0, k)$ , then  $\alpha_j = \mathcal{L}_{LS}(\alpha_l) \in \mathcal{S}_P(0, k)$ .

To show  $\mathcal{L}_{LS}$  is continuous, define the matrix  $R_p$  as the  $M$  by  $M - 2P$  matrix with

$$[R_p]_{i,j} = \begin{cases} 1 & \text{for } i = j + p \\ 1 & \text{for } i = 2P - p + j \\ 0 & \text{otherwise} \end{cases} \quad (3.71)$$

for  $i = 1 \dots M$  and  $j = 1 \dots M - 2P$ . Thus, for  $p = 0 \dots P$ ,  $R_p$  is of rank  $M - 2P$  and so is the sum of any  $P_o$  such matrices. The matrix  $A$  can now be decomposed as the sum of these matrices as

$$A = R_o + \sum_{p=1}^P \alpha[p] R_p = R_o + R(\alpha) \quad (3.72)$$

Using this decomposition, the  $M - 2P$  square matrix  $A^T K_w A$  can be written as a rational function of  $\alpha$  as

$$A^T K_w A = [R_o + R(\alpha)]^T K_w [R_o + R(\alpha)] \quad (3.73)$$

Because sums, products and inverses of rational functions are also rational functions, then  $\mathcal{L}_{LS}(\alpha)$  is a rational function of  $\alpha$  (70). Therefore,  $\mathcal{L}_{LS}$  is continuous everywhere it is defined. Now,  $A^T K_w A$  is invertible and of rank  $M - 2P$ . Since  $Y_c$  is assumed to be of rank  $P$ , then the matrix product  $Y_c^T [A^T K_w A]^{-1} Y_c$  is of rank  $P$  so that its inverse exists. Thus  $\mathcal{L}_{LS}(\alpha)$  is defined for all  $\alpha$  and, as such, is continuous on  $\mathcal{S}_P(0, k)$  (48:857).

Finally, to show  $\mathcal{L}_{LS}$  is bounded, the limit as  $\alpha$  grows unbounded must exist and be bounded (70). First, define the set

$$\mathcal{S}_P(0, 1) = \{\mathbf{x} \in R^P : \|\mathbf{x}\| = 1\} \quad (3.74)$$

and let  $\alpha = \beta \mathbf{x}$  for  $\beta \in R$ . Then

$$\begin{aligned} \mathcal{L}_{LS}(\beta \mathbf{x}) = & - \left\{ Y_c^T \left[ \left( \frac{R_o^T}{\beta} + R(\mathbf{x})^T \right) K_w \left( \frac{R_o}{\beta} + R(\mathbf{x}) \right) \right]^{-1} Y_c \right\}^{-1} \\ & \bullet Y_c^T \left[ \left( \frac{R_o^T}{\beta} + R(\mathbf{x})^T \right) K_w \left( \frac{R_o}{\beta} + R(\mathbf{x}) \right) \right]^{-1} \mathbf{y}_c \end{aligned} \quad (3.75)$$

Taking the limit as  $\beta$  approaches infinity yields

$$\lim_{\beta \rightarrow \infty} \mathcal{L}_{LS}(\beta \mathbf{x}) = - (Y_c^T [R^T(\mathbf{x}) K_w R(\mathbf{x})]^{-1} Y_c)^{-1} Y_c^T [R^T(\mathbf{x}) K_w R(\mathbf{x})]^{-1} \mathbf{y}_c = \mathcal{L}_{LS, \infty}(\mathbf{x}) \quad (3.76)$$

Since  $R(\mathbf{x})$  is still of rank  $M - 2P$ , then  $\mathcal{L}_{LS, \infty}(\beta \mathbf{x})$  is simply the solution to the weighted least squares problem and is therefore a finite vector with  $\|\mathcal{L}_{LS, \infty}(\beta \mathbf{x})\|^2 < \infty$ . Finally, let

$$f(\beta) = \max_{\mathbf{x} \in \mathcal{S}_P(0, 1)} \mathcal{L}_{LS}(\beta \mathbf{x}) \quad (3.77)$$

As shown in (70),  $f(\beta)$  is a continuous function with

$$\lim_{\beta \rightarrow \infty} f(\beta) = \max_{\mathbf{x} \in \mathcal{S}_P(0, k)} \mathcal{L}_{LS, \infty}(\mathbf{x}) < \infty \quad (3.78)$$

Since  $f(\beta)$  is a continuous function bounded at infinity, so is its norm and the supremum of  $f(\beta)$  is finite. That is (70)

$$\sup_{\alpha} \|\mathcal{L}_{LS}(\alpha)\| = \sup_{\alpha} \|f(\beta)\| < \infty \quad (3.79)$$

Thus, the supremum of the infinity norm of  $\mathcal{L}_{LS}$  is finite and maps the bounded set  $S_P(0, k) \subset R^P$  into itself. Finally, since  $\mathcal{L}_{LS}$  is a continuous function that maps a convex compact set into a convex compact set, then via Schauder's Theorem (33:152-153), there exists a fixed point such that  $\alpha \in S_P(0, k)$  such that  $\mathcal{L}_{LS}(\alpha) = \alpha$ . The fixed point obtained using an iterative form of the LS function of Equation 3.65 will become the estimate of the LP coefficients.

**3.5.3.3 Application.** From the true system model, where  $\alpha$  indicates the true LP coefficients,

$$y_c(\alpha_o) = -Y_c(\alpha_o)\alpha_o + e(\alpha_o) \quad (3.80)$$

Clearly, the quantity  $-y_c(\alpha_o) + e(\alpha_o)$  is in the range of  $Y_c(\alpha_o)$ . If the least squares estimate,  $\hat{\alpha}_o$  is given by Equation 3.65, substitution yields

$$\hat{\alpha}_o = [Y_c^T(\alpha_o)Y_c(\alpha_o)]^{-1} Y_c^T(\alpha_o) [Y_c(\alpha_o)\alpha_o - e(\alpha_o)] \quad (3.81)$$

For  $\hat{\alpha}_o$  to be equal to  $\alpha_o$ , the error vector must be orthogonal to the columns of  $Y_c(\alpha_o)$ . In this case, the true values of the LP coefficients,  $\hat{\alpha}_o$ , would be a fixed point of the Least Squares mapping

$$\alpha_o = \mathcal{L}_{LS}(\alpha_o) \quad (3.82)$$

Now consider the iterative form of the LS mapping function  $\mathcal{L}_{ILS}$ . From the system model, at iteration  $i$ ,

$$-y_c(\alpha_o^i) + e(\alpha_o^i) = Y_c(\alpha_o^i)\alpha_o^i \quad (3.83)$$

The Iterative Least Squares (ILS) estimate for the LP coefficients is found from Equation 3.65 as

$$\alpha_o^{i+1} = -[Y_c^T(\alpha_o^i)Y_c(\alpha_o^i)]^{-1} Y_c^T(\alpha_o^i)y(\alpha_o^i) = \mathcal{L}_{ILS}(\alpha_o^i) \quad (3.84)$$

At each iteration, the squared error, calculated as

$$J(\alpha_o^i) = e^T(\alpha_o^i)e(\alpha_o^i) \quad (3.85)$$

is minimized with respect to the current LP coefficient estimate. When  $\alpha_o^{i+1} = \alpha_o^i$ , the error vector,  $e(\alpha^i)$  is orthogonal to the columns of  $Y_c(\alpha^i)$  and  $\alpha_o^i = \mathcal{L}_{ILS}(\alpha_o^i)$ . A fixed point has been reached and the estimate of the LP coefficients becomes  $\hat{\alpha}_o = \alpha_o^i$ .

**3.5.3.4 Simulations.** Figure 3.7 shows a typical convergence of the ILS algorithm for one and two filtered sinusoids. In both cases, the center frequency of the filter was used to provide the initial LP estimates and less than ten iterations were required for location of the fixed point. In addition, the fixed point located does not coincide with the LP coefficients given by the minimum of the error surface. Figure 3.8 shows the estimation accuracy of the ILS fixed point mapping for a single sinusoid as a function of SNR and frequency. Figure 3.8a shows the ILS algorithm produces accurate frequency estimates only when the LP general linear model incorporates the effects of the filter upon the noise. Assumptions of uncorrelated (white) noise produce an incorrect model. In addition, provided the correct model is used, the accuracy of the estimator is relatively independent of the method of initialization above the SNR of about 3dB. For lower SNRs, the algorithm becomes sensitive to the initial LP estimates and the threshold can be extended to about 0dB using the filter center frequency to provide initial estimates. Figure 3.8b, shows the frequency estimation accuracy of the ILS fixed point estimator across the entire range of frequencies for a sinusoid in zero-mean, normally distributed, uncorrelated noise. Using the simple LS estimate as the initial estimate, this figure shows the ILS fixed point algorithm provides optimal frequency estimates except for extremely low ( $f < .01$ ) or high ( $f > .49$ ) frequencies.

Figure 3.9a shows the accuracy of the ILS fixed point estimator for the two widely spaced sinusoids in zero-mean, independent, normally distributed noise. Using an initial estimate obtained from the general LS solution, this figure shows the frequency estimates obtained as a result of the ILS estimator are near optimal for SNRs above threshold. As with a single sinusoid, Figure 3.9b shows the ILS algorithm produces accurate frequency

estimates only when the LP general linear model incorporates the effects of the filter upon the noise. Assumptions of uncorrelated noise produce an incorrect model and, consequently, suboptimal estimates. For two sinusoids however, initializing the algorithm with LP coefficients obtained near the center filter center frequency produced accurate estimates at a slightly lower SNRs than those produced by a simple LS initial estimate.

The ability to estimate the frequencies of both widely and closely separated sinusoids, either in independent or correlated, normally distributed noise, imply the ILS estimator can be used as a general frequency estimator.

### 3.5.4 Iterative Total Least Squares Fixed Point Mapping.

**3.5.4.1 Theory.** In forming the LS solution, the noise in the observation matrix,  $Y_c(\alpha_o)$ , was ignored. The LS solution simply found the vector of minimum norm which must be added to the vector  $-\mathbf{y}_c(\alpha_o)$  to bring it into the range of  $Y_c(\alpha_o)$ . The TLS solution attempts to account for the noise in the observation matrix,  $Y_c(\alpha_o)$  in addition to the noise in the  $\mathbf{y}_c(\alpha_o)$  (21),(67),(87:533-538). First, Equation 3.62 can be rewritten as

$$[\mathbf{y}_c(\alpha)|Y_c(\alpha)] \begin{bmatrix} 1 \\ \alpha_o \end{bmatrix} = 0 = Y_C(\alpha)\alpha \quad (3.86)$$

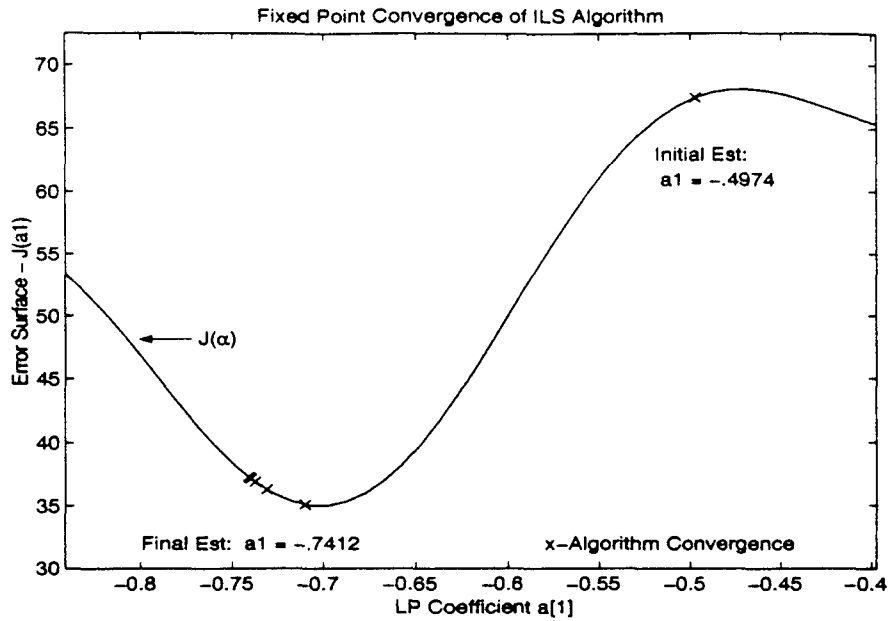
The TLS solution for  $\alpha$  seeks the matrix,  $E(\alpha)$ , of minimum Frobenius norm such that the matrix sum,  $[Y_C(\alpha) + E(\alpha)]$ , is rank deficient (21), (67). Decomposing  $Y_C(\alpha)$  via the SVD yields

$$Y_C(\alpha) = U(\alpha)\Sigma(\alpha)V^T(\alpha) \quad (3.87)$$

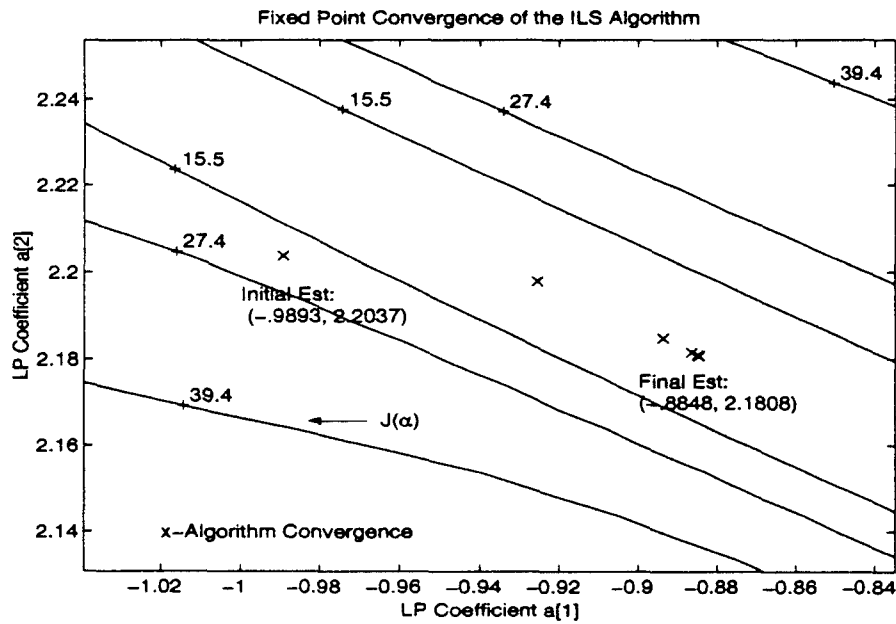
Here,  $U(\alpha)$  and  $V(\alpha)$  are unitary matrices which can be written in terms of their column vectors as

$$U(\alpha) = [\mathbf{u}_1(\alpha) \dots \mathbf{u}_{M-2P}(\alpha)] \quad \text{with } \mathbf{u}_i \in R^{M-2P} \quad (3.88)$$

$$V(\alpha) = [\mathbf{v}_1(\alpha) \dots \mathbf{v}_{P+1}(\alpha)] \quad \text{with } \mathbf{v}_i \in R^{P+1} \quad (3.89)$$

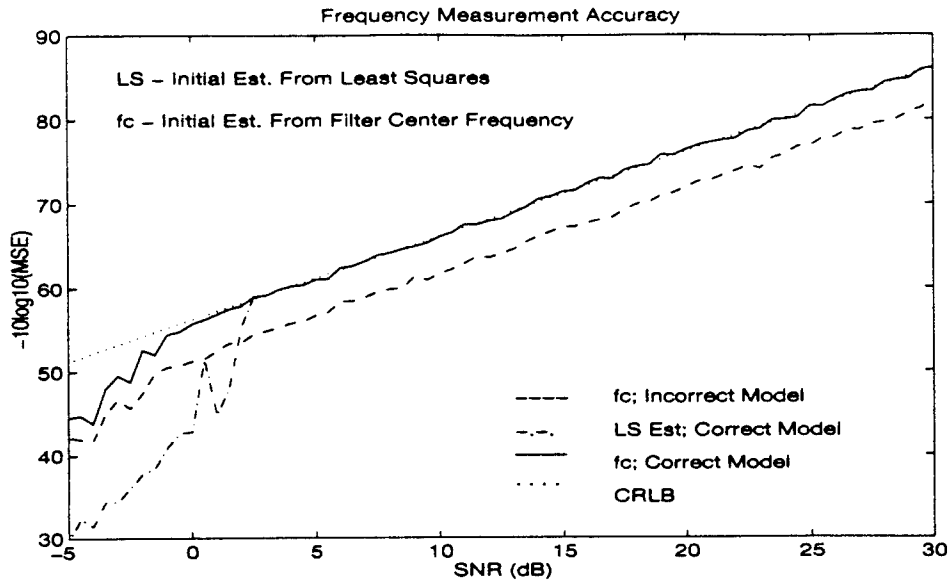


(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$  and  $SNR = -5\text{dB}$ .  
Final Objective Function Evaluation:  $J(\alpha_{ILS}) = 37.26$ ,  $J_{MIN} = 35.009$ .

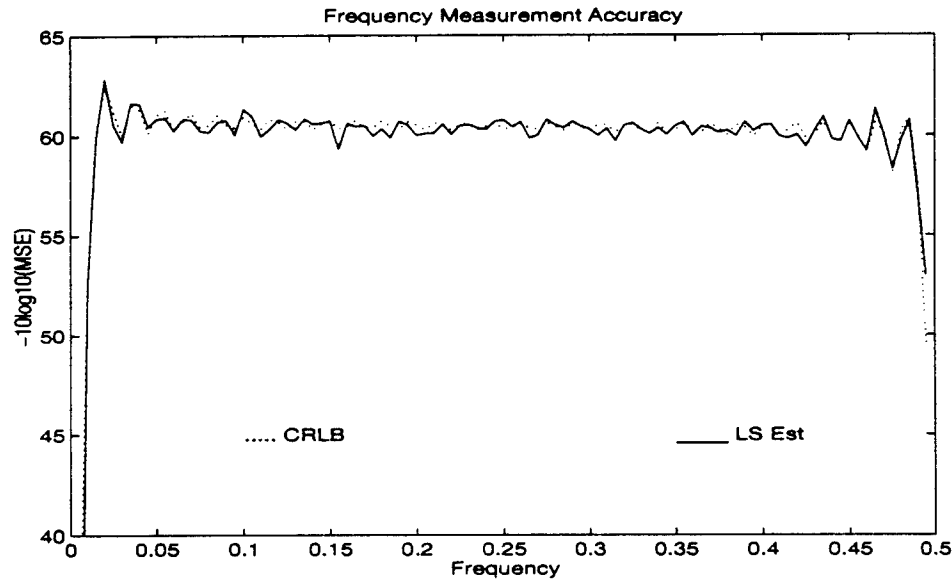


(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3]$ ,  $[b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .  
Objective Function Evaluation:  $J(\alpha_{ILS}) = 3.6088$ ,  $J_{MIN} = 3.5230$ ;  $SNR = 5\text{ dB}$ .

Figure 3.7 **ILS Fixed Point Convergence:**  $J(\alpha_o)$  calculated at fixed SNR and block length  $M = 32$ . Ten iterations allowed for convergence. Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ . SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ .

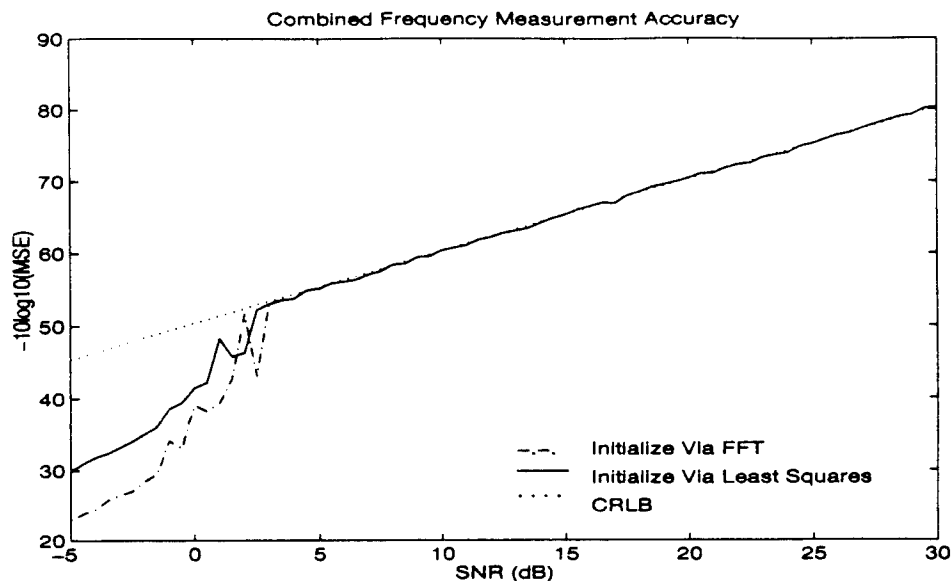


(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ , MSE at .5dB SNR intervals.  
Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ .

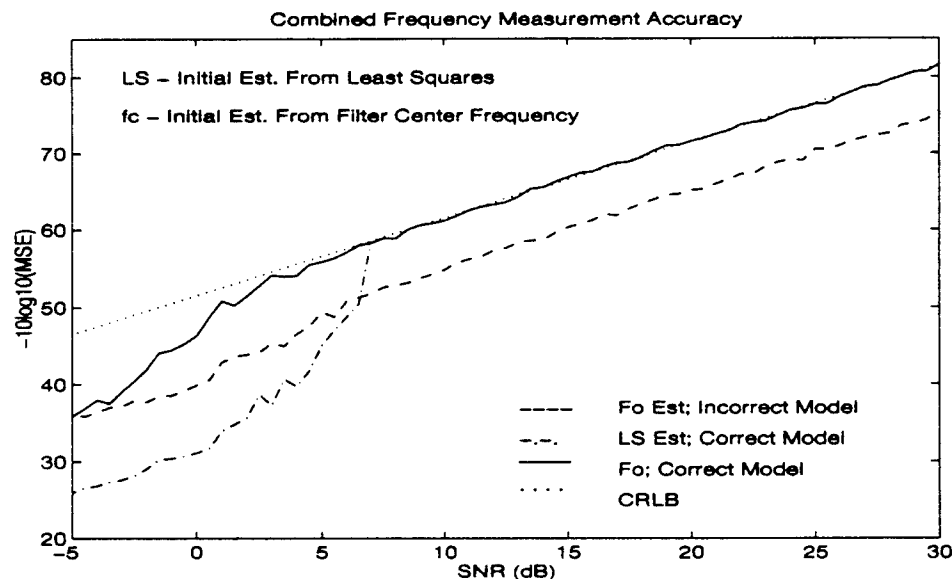


(b) Sinusoid Parameters:  $[b_1^o = 1, \phi_1^o = \pi/3]$ ,  $SNR = 10\text{dB}$ .  
MSE at .005 frequency intervals.

**Figure 3.8 ILS Frequency Estimation Accuracy-One Sinusoid:** MSE calculated from 500 independent realizations of block length  $M = 32$ . Ten iterations allowed for convergence. The incorrect model assumed  $K_w = \sigma^2 I$  while the correct model used  $K_w = \sigma^2 H^T H$ . SNR calculated as  $SNR = -10\log_{10}(2\sigma^2)$ ;  $MSE = \frac{1}{500} \sum_{i=1}^{500} (\hat{f}_1[i] - f_1)$ .



(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .1797, \phi_1^o = \pi/3]$ ,  $[b_2^o = 1, f_2 = .3828, \phi_2^o = 4\pi/3]$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3]$ ,  $[b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .  
Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$  Length:  $N = 32$ .

Figure 3.9 **ILS Frequency Estimation Accuracy-Two Sinusoids:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. The incorrect model assumed  $K_w = \sigma^2 I$  while the correct model used  $K_w = \sigma^2 H^T H$ . SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ .



In addition,  $\Sigma(\alpha)$  is an  $M - 2P$  by  $P + 1$  diagonal matrix of the singular values of  $Y_C(\alpha)$ , denoted  $\sigma_i$  for  $i = 1 \dots P + 1$ , arranged so that  $\sigma_1 \geq \sigma_2 \dots \geq \sigma_{P+1} > 0$ . The TLS solution for  $\alpha$  is given by the mapping function,  $\mathcal{L}_{TLS}$ , as (86:534)

$$\hat{\alpha} = \frac{v_{P+1}(\alpha)}{V(\alpha)_{1,P+1}} = \mathcal{L}_{TLS}(\alpha) \quad (3.90)$$

Provided  $V(\alpha)_{1,P+1} \neq 0$ , this solution is unique and provides the set of LP coefficients associated with the matrix of minimum Frobenius norm,  $E(\alpha)$  such that (21)

$$-y_c(\alpha) + e(\alpha) \in \text{Range}\{Y_c(\alpha) + E(\alpha)\} \quad (3.91)$$

In order to use this estimation technique for the LP coefficients, a fixed point must exist.

**3.5.4.2 Existence.** To show a fixed point exists for the TLS mapping function, let  $\mathcal{S}_{P+1}(0,1) : \{\alpha \in R^{P+1} | \alpha^T \alpha = 1\}$  and let  $\alpha_l \in \mathcal{S}_{P+1}(0,1)$ . The solution provided by  $\mathcal{L}_{TLS}(\alpha_l)$  is the right-most singular vector

$$\alpha_j = \frac{v_{P+1}(\alpha_l)}{V(\alpha_l)_{1,P+1}} \quad (3.92)$$

Here,  $V(\alpha_l)_{1,P+1}$  is the first element of the  $P + 1$  column of  $V(\alpha_l)$ . Since this solution can be scaled by any constant without changing the zeros of the LP polynomial  $A(z)$ , a modified form of the TLS mapping becomes

$$\alpha_j = v_{P+1}(\alpha_l) = \mathcal{L}_{TLS}(\alpha_l) \quad (3.93)$$

Since  $v_{P+1}(\alpha_l)^T v_{P+1}(\alpha_l) = 1$ , then  $\alpha_j = \mathcal{L}_{TLS}(\alpha_l) \in \mathcal{S}_{P+1}(0,1)$ .

Now, to show  $\mathcal{L}_{TLS}$  is continuous, the TLS solution was found via the SVD of the matrix product  $G(\alpha)Y_C$ . Since  $G(\alpha)$  is a square  $M - 2P$  nonsingular matrix and  $Y_C$  is and  $M - 2P$  by  $P + 1$  matrix of rank  $P + 1$ , then  $v_{P+1}(\alpha)$  exists for all  $\alpha$ . Furthermore, this solution is the same solution as the TLS solution of

$$Y_C^T [A_i^T K_w A_i]^{-1} Y_C \alpha = 0 \quad (3.94)$$

This function is a rational function of  $\alpha$ , and since sums, products and inverses of rational functions are also rational functions, then  $\mathcal{L}_{TLS}(\alpha)$  is continuous everywhere in its domain (48:857). Since the TLS solution exists for all  $\alpha \in \mathcal{S}_{P+1}(0,1)$ , then  $\mathcal{L}_{TLS}(\alpha)$  is continuous on  $\mathcal{S}_{P+1}(0,1)$ . Thus, via Brouwer's Fixed Point theorem, there exists an  $\alpha \in \mathcal{S}_{P+1}(0,1)$  such that a fixed point exists (33:116). The fixed point obtained via an iterative form of the TLS solution will become the estimate of the LP coefficients.

**3.5.4.3 Application.** From the true system model, where  $\alpha$  indicates the true LP coefficients

$$Y_c(\alpha)\alpha_o = -y_c(\alpha) + e(\alpha) \quad (3.95)$$

Again, the vector  $-y_c(\alpha) + e(\alpha)$  is in the range of  $Y_c(\alpha)$ . Rearranging terms gives

$$[Y_c(\alpha) - W_c(\alpha)]\alpha = 0 \quad (3.96)$$

Decomposing  $Y_c(\alpha)$  as the SVD yields the TLS estimate as

$$\hat{\alpha} = \frac{v_{P+1}(\alpha)}{V(\alpha)_{1,P+1}} \quad (3.97)$$

For  $\hat{\alpha}$  to be equal to  $\alpha$ , the matrix,  $-W_c(\alpha)$ , must be of minimum Frobenius norm. In this case, the true values of the LP coefficients become a fixed point of the TLS function

$$\alpha = \mathcal{L}_{TLS}(\alpha) \quad (3.98)$$

Now consider the iterative form of the TLS solution. From the system model, at iteration  $i$ , the model yields

$$Y_c(\alpha^i)\alpha^i = e(\alpha^i) \quad (3.99)$$

The Iterative Total Least Squares (ITLS) estimate for the LP coefficients is found via the mapping function,  $\mathcal{L}_{ITLS}$ , as

$$\alpha^{i+1} = \frac{v_{P+1}(\alpha^i)}{V(\alpha^i)_{1,P+1}} = \mathcal{L}_{ITLS}(\alpha^i) \quad (3.100)$$

At each iteration, the squared error is calculated as

$$J(\alpha) = e^T(\alpha^i)e(\alpha^i) \quad (3.101)$$

When  $\alpha^{i+1} = \alpha^i$ , the noise matrix,  $W(\alpha^i)$  is of minimum Frobenius norm and the fixed point has been reached so that  $\alpha^i = \mathcal{L}_{ITLS}(\alpha^i)$ .

**3.5.4.4 Simulations.** Figure 3.10 shows a typical convergence of the ITLS algorithm for one and two filtered sinusoids. In both cases, the center frequency of the filter was used to provide the initial LP estimates and less than ten iterations were required for location of the fixed point. In addition, the fixed point located does not coincide with the LP coefficients given by the minimum of the error surface. Figure 3.11 shows the estimation accuracy of the ITLS fixed point mapping for a single sinusoid as a function of SNR and frequency. Figure 3.11a shows the ITLS algorithm produces accurate frequency estimates only when the LP general linear model incorporates the effects of the filter upon the noise. Assumptions of uncorrelated (white) noise produce an incorrect model. In addition, provided the correct model is used, the accuracy of the estimator is relatively independent of the method of initialization above the SNR of about 3dB. For lower SNRs, the algorithm becomes sensitive to the initial LP estimates and the threshold can be extended to about 0dB using the filter center frequency to provide initial estimates. Figure 3.11b, shows the frequency estimation accuracy of the ITLS fixed point estimator across the entire range of frequencies for a single sinusoid in zero mean, normally distributed, uncorrelated noise. Using the simple LS estimate as the initial estimate, this figure shows the ITLS fixed point algorithm can provide optimal frequency estimates except for extremely low ( $f < .01$ ) or high ( $f > .49$ ). frequencies.

Figure 3.12a shows the accuracy of the ITLS fixed point estimator for the two widely-spaced sinusoids in zero-mean, independent, normally distributed noise. Using an initial estimate obtained from the general TLS solution, this figure shows the frequency estimates obtained as a result of the ITLS estimator are near optimal for SNRs above threshold.

As for a single sinusoid, Figure 3.12b shows the ITLS algorithm produces accurate frequency estimates only when the LP general linear model incorporates the effects of the

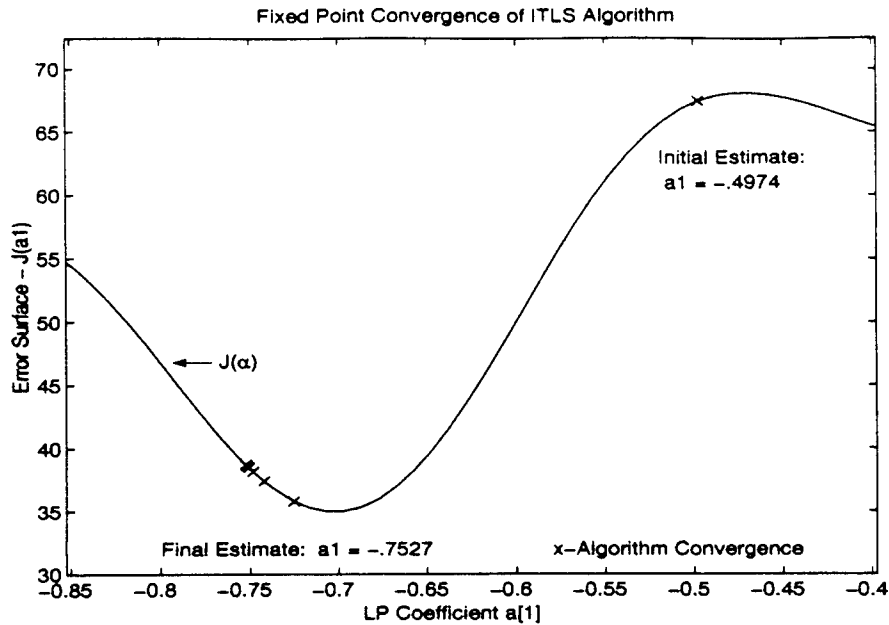
filter upon the noise. Assumptions of uncorrelated noise produce an incorrect model and, consequently, suboptimal estimates. For two sinusoids however, initializing the algorithm with LP coefficients obtained near the filter center frequency produced accurate estimates at a slightly lower SNRs than those produced by a simple TLS initial estimate.

The ability to estimate the frequencies of both widely and closely separated sinusoids, either in independent or correlated, normally distributed noise, imply the ITLS estimator can be used as a general frequency estimator. However, this estimator requires an SVD of the  $M - 2P$  by  $P + 1$  observation matrix at each iteration.

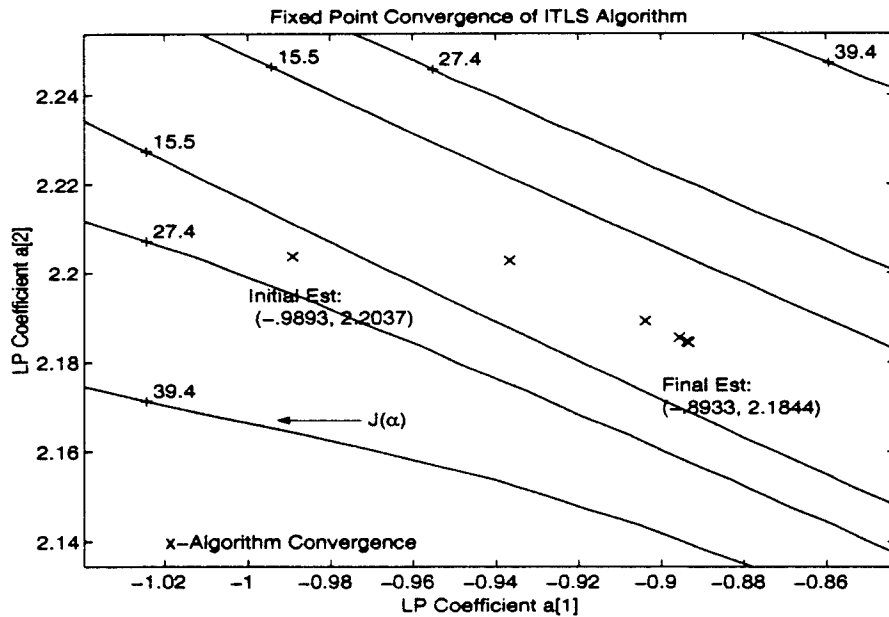
*3.5.5 Section Summary.* In this section, two methods, based on fixed point theory, were derived for estimating the coefficients of the LP general linear model. One method, termed the ILS algorithm, is based to an iterative least squares solution on an over-determined system of equations while the other method, coined the ITLS algorithm, is based on an iterative total least squares solution. Simulations indicate both methods provide minimum variance unbiased estimates of the LP coefficients, and consequently, the sinusoidal frequencies, over a wide range of SNRs. All other factors being equal, the ILS method would be preferred over the ITLS method since the ITLS algorithm requires an SVD of a  $M - 2P$  by  $P + 1$  matrix at each iteration. Application of fixed point theory to the estimation of the LP coefficients and development of the ILS and ITLS fixed point mapping functions is an original contribution of this research. The next section relates the estimates found by these methods to the ML estimates of the LP coefficients.

### *3.6 LP Modeling/ ML Frequency Estimation Relationship*

This section derives the exact relationship between the true LP general linear model and ML Frequency estimation. Specifically, this section will show the LP coefficients, defined by the general linear model of Equation 3.53, will provide ML frequency estimates for filtered data provided the two-norm of the model error is minimized. This derivation, establishing an exact relationship between ML frequency estimation of filtered sinusoids and the LP general linear model, is an original contribution of this dissertation.

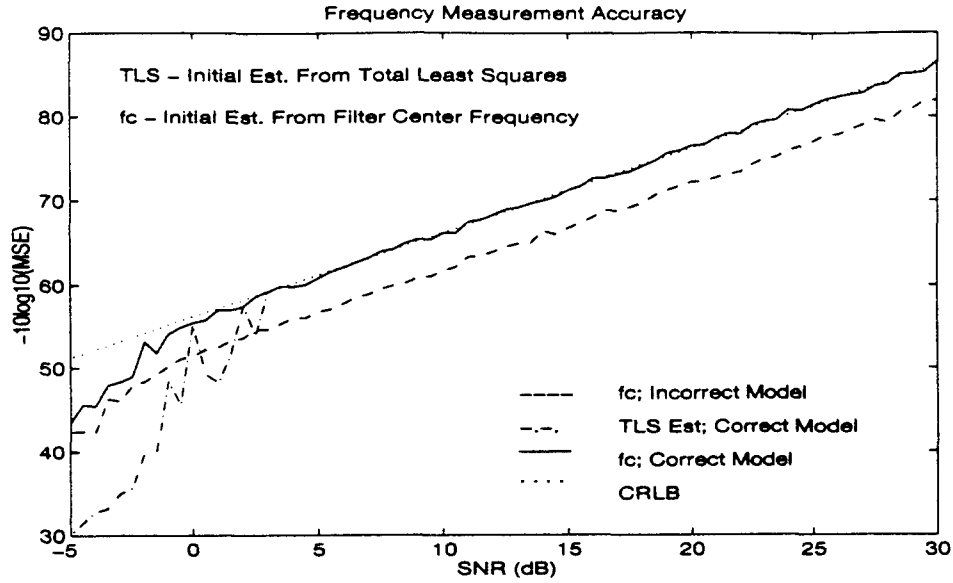


(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ , and  $SNR = -5dB$ .  
Objective Function Evaluation:  $J(\alpha_{ITLS}) = 38.71$ ,  $J_{MIN} = 35.009$ .

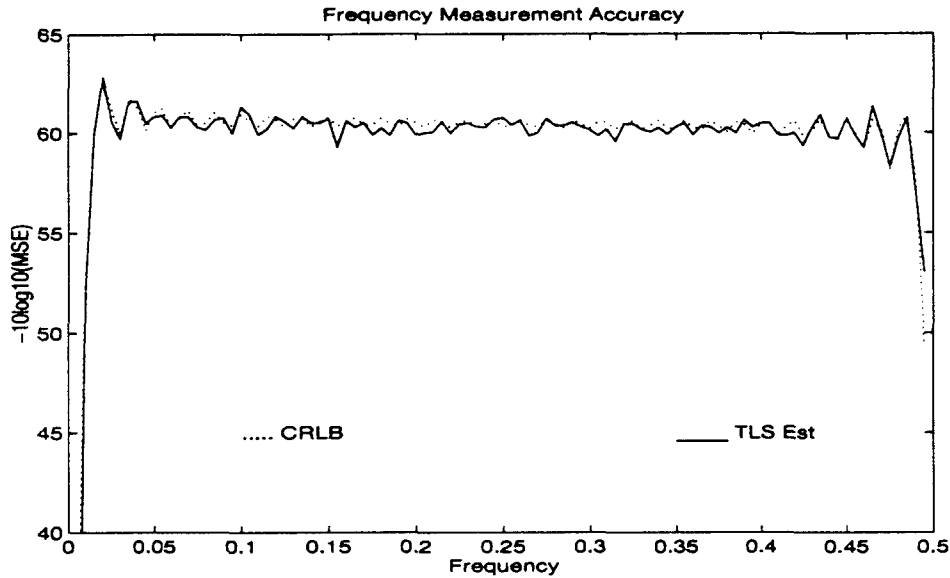


(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3]$ ,  $[b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .  
Objective Function Evaluation:  $J(\alpha_{ITLS}) = 3.6883$ ,  $J_{MIN} = 3.5230$ ;  $SNR = 5dB$ .

Figure 3.10 **ITLS Fixed Point Convergence:**  $J(\alpha)$  calculated as function of LP coefficients at fixed SNR and block length  $M = 32$ . Ten iterations allowed for convergence. Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ . SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ .

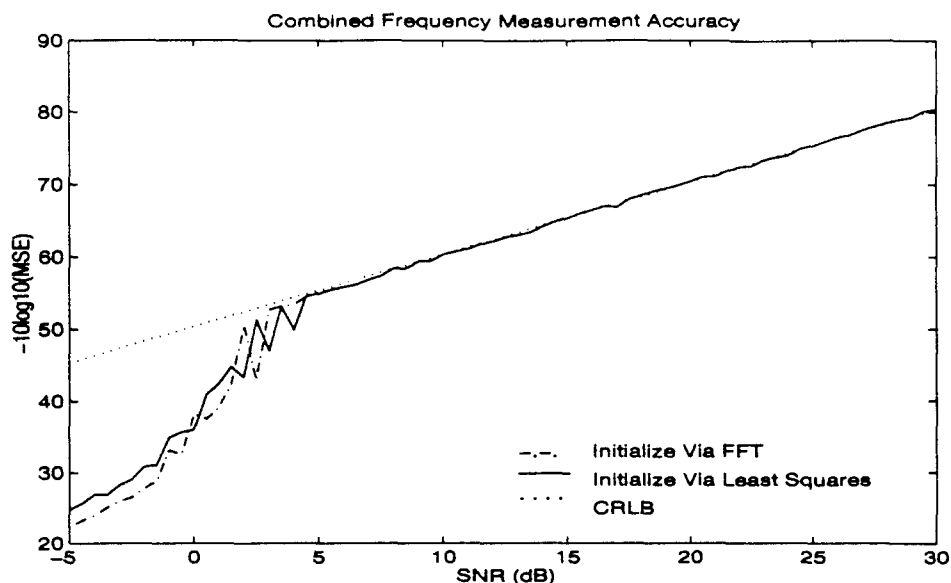


(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ , MSE at .5dB SNR intervals.  
Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ .

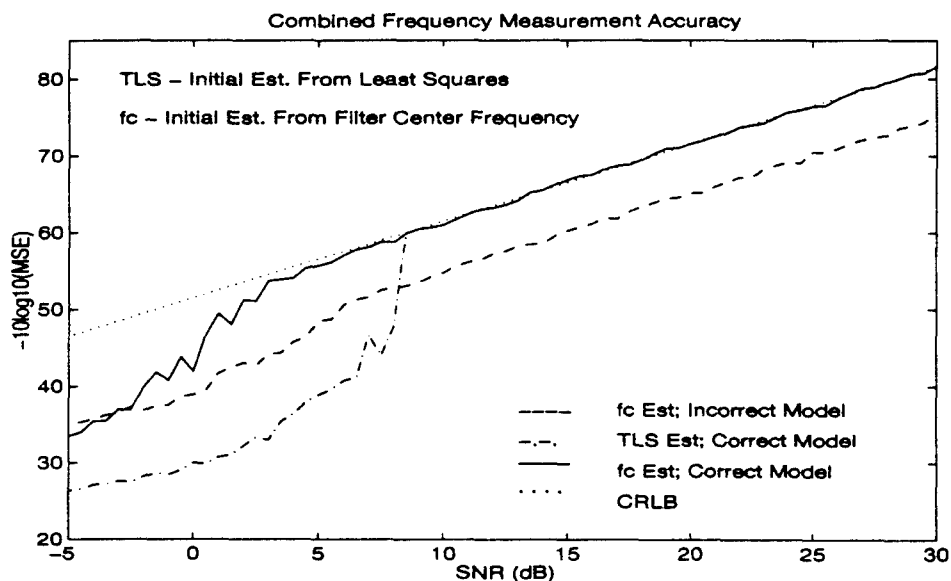


(b) Sinusoid Parameters:  $[b_1^o = 1, \phi_1^o = \pi/3]$ ,  $SNR = 10\text{dB}$ .  
MSE at .005 frequency intervals.

Figure 3.11 **ITLS Frequency Estimation Accuracy-One Sinusoid:** MSE calculated from 500 independent realizations of block length  $M = 32$ . Ten iterations allowed for convergence. The incorrect model assumed  $K_w = \sigma^2 I$  while the correct model used  $K_w = \sigma^2 H^T H$ . SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ ;  $MSE = \frac{1}{500} \sum_{i=1}^{500} (\hat{f}_1[i] - f_1)$ .



(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .1797, \phi_1^o = \pi/3]$ ,  $[b_2^o = 1, f_2 = .3828, \phi_2^o = 4\pi/3]$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3]$ ,  $[b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .  
Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ .

Figure 3.12 ITLS Frequency Estimation Accuracy-Two Sinusoids: MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. The incorrect model assumed  $K_w = \sigma^2 I$  while the correct model used  $K_w = \sigma^2 H^T H$ . SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^P \sum_{i=1}^{500} (\hat{f}_p(i) - f_p)^2$ .

3.6.1 *ML Objective Function Redefined.* From Equation 3.25, to provide ML frequency estimates, the objective function  $J(\mathbf{f})$  must be minimized with respect to  $\mathbf{f}$ .

$$J(\mathbf{f}) = \mathbf{y}^T K_{\mathbf{w}}^{-1} \mathbf{y} - \mathbf{y}^T K_{\mathbf{w}}^{-1} \Lambda [\Lambda^T K_{\mathbf{w}}^{-1} \Lambda]^{-1} \Lambda^T K_{\mathbf{w}}^{-1} \mathbf{y}$$

To minimize this ML frequency objective function, consider a vector space interpretation of the minimization problem and assume that  $\mathbf{y}$  has length  $M > 2P$ . From Equation 3.13, since  $\Lambda$  is an  $M$  by  $2P$  matrix with  $2P$  linearly independent columns, the column vectors of  $\Lambda$ , denoted

$$\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_{2P}] \quad (3.102)$$

span a  $2P$  dimensional subspace,  $V_1$ , of the real vectors,  $R^M$ . As such, there exists an  $M - 2P$  dimensional orthogonal subspace, denoted  $V_2$ , which is spanned by the column vectors in the matrix  $A$ ,

$$A = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{M-2P}] \quad (3.103)$$

For any vector  $\psi \in V_1$  and  $\gamma \in V_2$  then  $\psi^T \gamma = 0$  (83:132-152). In particular,  $\lambda_i^T \mathbf{a}_j = 0$  for  $i = 1 \dots 2P$  and  $j = 1 \dots M - 2P$ . Now let  $F$  be an  $M$  by  $M$  invertible matrix and define the linear transformation upon the two sets of basis vectors as  $\psi_i = F \lambda_i$ , where  $\psi_i \in V_1$ , and  $\gamma_j = F^{-T} \mathbf{a}_j$ , where  $\gamma_j \in V_2$ . In matrix form,

$$\Psi = [\psi_1 \dots \psi_{2P}] = F[\lambda_1 \dots \lambda_{2P}] = F\Lambda \quad (3.104)$$

$$\Gamma = [\gamma_1 \dots \gamma_{M-2P}] = F^{-T}[\mathbf{a}_1 \dots \mathbf{a}_{M-2P}] = F^{-T}A \quad (3.105)$$

The columns of  $\Psi$  form a basis for  $V_1$  while the columns of  $\Gamma$  form a basis for  $V_2$  (26:92). The projection of any vector  $\mathbf{z} \in R^M$  into  $V_1$  and  $V_2$  can be written in terms of the matrices  $\Psi$  and  $\Gamma$  as (83:156),(73:46-47)

$$\mathbf{z}_1 = \Psi [\Psi^T \Psi]^{-1} \Psi^T \mathbf{z}; \quad \text{for } \mathbf{z}_1 \in V_1 \quad (3.106)$$

$$\mathbf{z}_2 = \Gamma [\Gamma^T \Gamma]^{-1} \Gamma^T \mathbf{z}; \quad \text{for } \mathbf{z}_2 \in V_2 \quad (3.107)$$



The signal energy,  $E$ , defined by the vector two norm  $\|z\|^2$  for  $z \in R^M$ , is projected orthogonally into each subspace (26:181-226) as  $E = z^T z = E_1 + E_2$  where

$$E_1 = z_1^T z_1 = z^T \Psi [\Psi^T \Psi]^{-1} \Psi^T z = z^T [I - \Gamma(\Gamma^T \Gamma)^{-1} \Gamma^T] z \quad (3.108)$$

$$E_2 = z_2^T z_2 = z^T \Gamma [\Gamma^T \Gamma]^{-1} \Gamma^T z = z^T [I - \Psi(\Psi^T \Psi)^{-1} \Psi^T] z \quad (3.109)$$

For a fixed energy  $E$ , minimizing  $E_1$  maximizes  $E_2$  and vice versa. Now let  $F$  be defined by the Cholesky decomposition of the noise covariance matrix so that

$$K_w^{-1} = F^T F \quad (3.110)$$

The ML frequency objective function,  $J(f)$ , can then be written as

$$J(f) = y^T F^T F y - y^T F^T F \Lambda (\Lambda^T F^T F \Lambda)^{-1} \Lambda^T F^T F y \quad (3.111)$$

For  $z = Fy$ , and  $\Psi = F\Lambda$ , an alternate form of  $J(f)$  becomes

$$J(f) = z^T z - z^T \Psi [\Psi^T \Psi]^{-1} \Psi^T z \quad (3.112)$$

Factoring the vector  $z$  and employing Equation 3.109 yields

$$J(f) = z^T [I - \Psi(\Psi^T \Psi)^{-1} \Psi^T] z = z^T \Gamma(\Gamma^T \Gamma)^{-1} \Gamma^T z \quad (3.113)$$

Minimizing the ML frequency objective function,  $J(f)$ , is equivalent to minimizing  $E_2$ , the energy of the signal projected into  $V_2$ . From Equation 3.109, since  $\Gamma = F^{-T} A$ , this minimization can be written in terms of the basis vectors  $a_j$  in  $A$  as

$$J(A) = y^T F^T F^{-T} A (A^T F^{-1} F^{-T} A)^{-1} A^T F^{-1} F y = y^T A (A^T K_w A)^{-1} A^T y \quad (3.114)$$

Minimizing the ML frequency objective function defined by Equation 3.25 with respect to  $f$  has been recast as minimizing Equation 3.114 with respect to the basis vectors in

the matrix  $A$ . Proper selection of these basis vectors will prove critical for obtaining a simplified estimator.

**3.6.2 Determination of the Basis Vectors.** To determine an optimal set of basis vectors,  $\mathbf{a}_k$ , consider the LP model of Equation 2.14. Substitution of  $s[m] = b_k \cos(\omega_k m + \phi_k)$  and expanding yields

$$\sum_{k=1}^P \sum_{p=0}^{2P} a[p] b_k \{ \cos(\omega_k [m-p]) \cos(\phi_k) - \sin(\omega_k [m-p]) \sin(\phi_k) \} = 0 \quad (3.115)$$

Since Equation 3.115 must hold for all  $\omega_k$  and  $2P \leq m \leq M-1$ , using the substitution,  $b_k^s = b_k \sin(\phi_k)$  and  $b_k^c = b_k \cos(\phi_k)$ , gives the property

$$\sum_{p=0}^{2P} a[p] \cos(\omega_k [m-p]) = 0 \quad (3.116)$$

$$\sum_{p=0}^{2P} a[p] \sin(\omega_k [m-p]) = 0 \quad (3.117)$$

Now define the vector of  $k$  zeros as  $\mathbf{0}_k$  and the  $M$ -dimensional vector  $\mathbf{a}_k$  as

$$\mathbf{a}_k^T = [\mathbf{0}_k^T, a[0] \dots a[2P], \mathbf{0}_{M-2P-k-1}^T] \quad (3.118)$$

for  $k = 0 \dots M - 2P - 1$ . Since the form of the basis vectors for  $V_1$  can be written as

$$\lambda_i = [\cos(2\pi f_i [M_o - 1]) \dots \cos(2\pi f_i [N - 1])]^T \quad (3.119)$$

$$\lambda_{P+i} = -[\sin(2\pi f_i [M_o - 1]) \dots \sin(2\pi f_i [N - 1])]^T \quad (3.120)$$

for  $i = 1 \dots P$ , then  $\mathbf{a}_k^T \lambda_j = 0$  for  $j = 1 \dots 2P$ . The  $M - 2P$  vectors  $\mathbf{a}_k$  are orthogonal to the columns of  $\Lambda$ . Furthermore, employing these vectors as the columns of  $A$  yields the LP coefficient matrix defined by Equation 3.44. The  $M - 2P$  columns of  $A$  are linearly independent and thus form a basis for  $V_2$ , the subspace orthogonal to  $V_1$ . In addition, each column of  $A$  is derived from a single vector  $\mathbf{a}$  so that  $J(A)$  can be written as

$$J(\mathbf{a}) = \mathbf{y}^T A (A^T K_w A)^{-1} A^T \mathbf{y} \quad (3.121)$$

Minimization of  $J(\mathbf{f})$  with respect to  $\mathbf{f}$  has been recast as a minimization of  $J(\mathbf{a})$  with respect to  $\mathbf{a}$ . The LP coefficients which minimize Equation 3.121 provide ML frequency estimates (6). Furthermore, via the invariance property of ML estimators, the  $\mathbf{a}$  which minimizes  $J(\mathbf{a})$  is the ML estimate of the LP coefficients(40:185). As shown in the next section, this ML estimate of the LP coefficients can also be obtained from the LP general linear model.

**3.6.3 Relationship to LP Model.** To show the relationship between ML frequency estimation and the LP general linear model, let  $G(\mathbf{a})$  be the square  $M-2P$  matrix obtained from the Cholesky decomposition of the matrix  $[A^T K_w A]^{-1}$ . The ML frequency objective function can now be written as

$$J(\mathbf{a}) = \mathbf{y}^T A G^T(\mathbf{a}) G(\mathbf{a}) A^T \mathbf{y} \quad (3.122)$$

Employing the  $M-2P$  by  $2P+1$  matrix of measurements,  $Y_F$ , and imposing the sinusoidal constraints, so that  $\mathbf{a} = B\boldsymbol{\alpha}$ , yields  $A^T \mathbf{y} = Y_F \mathbf{a} = Y_C \boldsymbol{\alpha}$ . Thus, an alternative form of Equation 3.121 becomes

$$J(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^T Y_C^T G(\boldsymbol{\alpha})^T G(\boldsymbol{\alpha}) Y_C \boldsymbol{\alpha} \quad (3.123)$$

By defining the error vector as  $\mathbf{e}(\boldsymbol{\alpha}) = [e[M-1] \dots e[2P]]^T$  and constraining  $\alpha[0] = 1$  yields

$$\mathbf{e}(\boldsymbol{\alpha}) = G(\boldsymbol{\alpha}) Y_C \boldsymbol{\alpha} = \mathbf{y}_e(\boldsymbol{\alpha}) + Y_c(\boldsymbol{\alpha}) \alpha_o \quad (3.124)$$

This equation is identical to the LP general linear model given by Equation 3.53. By properly incorporating the effects of the filter coefficients and the LP coefficients in developing the LP general linear model, this derivation shows the exact relationship between the LP coefficients and ML frequency estimates. The set of LP coefficients which minimize the squared error defined by the LP general linear model will provide ML estimates of the frequencies. Furthermore, via the invariance of ML estimators, the LP coefficients which minimize the squared error are the ML estimates of the coefficients.

*3.6.4 Section Summary.* This section derived the exact relationship between the ML frequency objective function for filtered sinusoids and the LP general linear model and is an original contribution of this research. Specifically, this section showed the set of LP coefficients which minimize the square error defined by the LP general linear model will provide ML frequency estimates and vice versa. The next section will derive an ML estimator for the LP coefficients and derive the relationship between this estimator and the ILS and ITLS fixed point estimators.

### *3.7 LP Coefficient Estimation: Objective Function Minimization*

In this section, after deriving a method, termed Iterative Exact Gradient Descent (IEGD), for obtaining exact ML frequency estimates using gradient descent algorithms to minimize the constrained form of the LP objective function,  $J(\alpha)$ , approximate ML estimation techniques are examined. Specifically, this section will show the LP coefficients found via the ILS fixed point mapping function are exactly equivalent to the Iterative Generalized Least Squares (IGLS) approximate ML estimator developed in (31),(32)(70). In addition, the LP coefficients found via the ITLS fixed point mapping function will be shown to be exactly equivalent to the well known Iterative Quadratic Maximum Likelihood (IQML) approximate ML estimator proposed in (6), (29), (30), (68), (63),(75). The proof of the equivalence of these estimators is an original contribution of this dissertation and correctly casts the IGLS and IQML estimators as fixed point mappings.

*3.7.1 Iterative Exact Gradient Descent (IEGD).* From the form of the objective function  $J(\alpha)$ , as defined in Equation 3.123, the interpretation as the the familiar squared error objective function is self evident

$$J(\alpha) = e(\alpha)^T e(\alpha) = \alpha^T Y_C^T G(\alpha)^T G(\alpha) Y_C \alpha \quad (3.125)$$

Minimizing  $J(\alpha)$  with respect to  $\alpha$  becomes a  $P + 1$  dimensional search for an estimate,  $\hat{\alpha}$ , such that

$$\hat{\alpha} = \arg\{\min_{\alpha} (\alpha^T Y_C^T G(\alpha)^T G(\alpha) Y_C \alpha)\} \quad (3.126)$$

The  $\alpha$  which minimizes  $J(\alpha)$  will provide ML estimates of the frequencies and the LP coefficients.

There are many methods for minimizing  $J(\alpha)$  with respect to  $\alpha$ . The method developed in this dissertation, termed the Iterative Exact Gradient Descent (IEGD) algorithm, is based on an iterative gradient descent algorithm of the general form (15:140-141), (40:187),

$$\alpha^{i+1} = \alpha^i - \left[ D^{-1} \frac{\partial J(\alpha)}{\partial \alpha} \right] \Big|_{\alpha=\alpha^i} \quad (3.127)$$

where the vector  $\partial J(\alpha)/\partial \alpha$  is the  $P + 1$  dimensional gradient vector defined as (40:187)

$$\left[ \frac{\partial J(\alpha)}{\partial \alpha} \right]_i = \frac{\partial J(\alpha)}{\partial \alpha[i]} \quad (3.128)$$

In addition,  $D$  is the Hessian of  $J(\alpha)$  with (40:187)

$$[D]_{i,j} = \frac{\partial^2 J(\alpha)}{\partial \alpha[i] \partial \alpha[j]} \quad (3.129)$$

As shown in Appendix E, the gradient is relatively simple to implement while the Hessian is a bit more complicated. Again, as with most iterative optimization techniques, there is no guarantee the algorithm will ever converge. Even if it does converge, there is no guarantee the global minimum has been attained. The solution achieved will only provide a local minimum and therefore the accuracy of the estimate will depend greatly on the location of the initial estimate.

Figure 3.13 shows typical solutions for the LP coefficients obtained using the IEGD algorithm. In this figure, the results of the ILS fixed point estimator were used as an initial guess for the IEGD algorithm. This figure shows the IEGD algorithm does indeed find the minimum of  $J(\alpha)$  and, consequently, the ML estimate of the LP coefficients and frequencies. The main drawback to the IEGD algorithm, when compared to the ILS and ITLS fixed point algorithms, is the added complexity of computing the Hessian at each iteration. Typical methods for reducing this complexity treat the matrix  $A^T K_w A$  as a constant, impose a constraint on  $\alpha$  and perform the minimization with respect to this

constraint (6),(9), (29), (31),(32),(75). The two estimators based on this technique have been termed the IGLS and IQML estimators.

**3.7.2 Iterative Generalized Least Squares (IGLS).** One method of performing an approximate minimization is to constrain  $\alpha[0] = 1$  and perform a gradient descent using Equation 3.127 with the matrix product,  $C(\alpha_o) = A^T K_w A$ , treated as a constant. For this constraint, the objective function becomes a function of the  $P$  remaining terms in  $\alpha_o$

$$J(\alpha_o) = \mathbf{y}_c^T [C(\alpha_o^i)]^{-1} \mathbf{y}_c + 2\mathbf{y}_c^T [C(\alpha_o^i)]^{-1} Y_c \alpha_o + \alpha_o^T Y_c^T [C(\alpha_o^i)]^{-1} Y_c \alpha_o \quad (3.130)$$

where  $C(\alpha_o^i)$  is derived from the previous estimate of  $\alpha_o$ . Taking the gradient and evaluating at  $\alpha_o^i$  yields the general iterative gradient descent algorithm

$$\alpha_o^{i+1} = \alpha_o^i - 2D^{-1} \{Y_c^T [C(\alpha_o^i)]^{-1} Y_c \alpha_o^i + Y_c^T [C(\alpha_o^i)]^{-1} \mathbf{y}_c\} \quad (3.131)$$

Employing the Hessian matrix,  $D$ , yields

$$D = 2Y_c^T [C(\alpha_o^i)]^{-1} Y_c \quad (3.132)$$

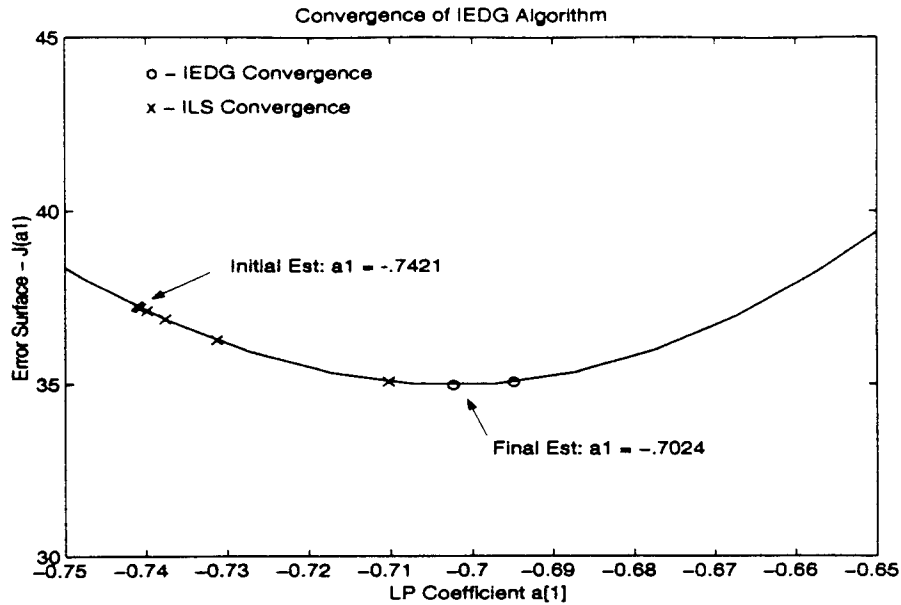
Thus, the final form of this iterative gradient descent algorithm becomes

$$\alpha_o^{i+1} = - \{Y_c^T [C(\alpha_o^i)]^{-1} Y_c\}^{-1} Y_c^T [C(\alpha_o^i)]^{-1} \mathbf{y}_c \quad (3.133)$$

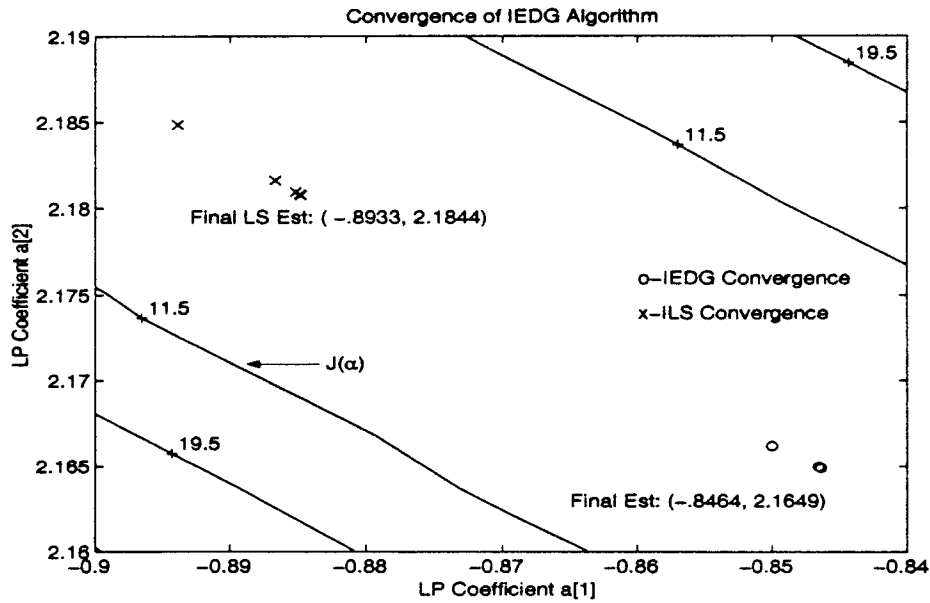
Since  $[C(\alpha_o^i)]^{-1} = G^T(\alpha_o^i)G(\alpha_o^i)$ , and  $Y_c(\alpha_o^i) = G(\alpha_o^i)Y_c$ , substitution yields

$$\alpha_o^{i+1} = - [Y_c^T(\alpha_o^i)Y_c(\alpha_o^i)]^{-1} Y_c^T(\alpha_o^i)\mathbf{y}_c(\alpha_o^i) \quad (3.134)$$

The approximate minimization of the ML objective function is identical to the ILS fixed point mapping function,  $\mathcal{E}_{ILS}(\alpha)$ , defined in Equation 3.84. This derivation is an original contribution of this dissertation for it establishes the connection between the IGLS algorithm developed in (31), (32) and (70) to the general method of gradient descent minimization and correctly casts the IGLS algorithm as a fixed point ILS mapping.



(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ , and  $SNR = -5dB$ .  
Objective Function Evaluation:  $J(\alpha_{IEGD}) = 35.009, J_{MIN} = 35.009$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3], [b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .  
Objective Function Evaluation:  $J(\alpha_{IEGD}) = 3.5230, J_{MIN} = 3.5230, SNR = 5dB$ .

Figure 3.13 **IEDG Algorithm Convergence:**  $J(\alpha)$  calculated as function of LP coefficients at fixed SNR and block length  $M = 32$ . Ten iterations allowed for convergence. Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ . SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ .

3.7.3 *Iterative Quadratic Maximum Likelihood (IQML)*. An alternative method of performing the approximate minimization of  $J(\alpha)$ , with  $C(\alpha) = A^T K_w A$  held constant, is to constrain  $\alpha$  to be of unit length,  $\alpha^T \alpha = 1$ , and iteratively minimize  $J(\alpha)$  via an eigenvalue analysis. That is, rewriting  $J(\alpha)$  as

$$J(\alpha) = \alpha^T Y_C^T [C(\alpha^i)]^{-1} Y_C \alpha = \alpha^T Y_C^T(\alpha^i) Y_C(\alpha^i) \alpha \quad (3.135)$$

Again,  $C(\alpha^i)$  represents a constant matrix obtained from the previous estimate of  $\alpha^i$ . Using the method of Lagrange Multipliers yields the objective function (49:897-923)

$$J_1(\alpha) = \alpha^T Y_C^T(\alpha^i) Y_C(\alpha^i) \alpha - \lambda \alpha^T \alpha \quad (3.136)$$

Taking the gradient with respect to  $\alpha$  and equating to zero produces

$$2Y_C^T(\alpha^i) Y_C(\alpha^i) \alpha - 2\lambda \alpha = 0 \quad (3.137)$$

Under this unit length constraint,  $\alpha$  is an eigenvector of the matrix product  $\{Y_C^T(\alpha^i) Y_C(\alpha^i)\}$ . Substituting into  $J(\alpha)$  gives

$$J(\alpha) = \lambda \alpha^T \alpha = \lambda \quad (3.138)$$

The optimal estimate of  $\alpha$  is the eigenvector associated with the smallest eigenvalue of the matrix product  $\{Y_C^T(\alpha^i) Y_C(\alpha^i)\}$ . Thus, for the IQML algorithm, each time an estimate of  $\alpha$  is obtained, it is reinserted into  $J(\alpha)$  to produce a new estimate. This process continues until convergence of the algorithm. An alternative explanation of this algorithm can be obtained by considering the SVD of  $Y_C(\alpha^i)$  directly. That is, let

$$Y_C(\alpha) = U(\alpha) \Sigma(\alpha) V^T(\alpha) \quad (3.139)$$

Here,  $U(\alpha)$  and  $V(\alpha)$  are unitary matrices. Writing in terms of column vectors yields

$$U(\alpha) = [u_1(\alpha) \dots u_{M-2P}(\alpha)] \quad \text{with } u_i \in R^{M-2P} \quad (3.140)$$

$$V(\alpha) = [v_1(\alpha) \dots v_{P+1}(\alpha)] \quad \text{with } v_i \in R^{P+1} \quad (3.141)$$



In addition,  $\Sigma(\alpha)$  is an  $M - 2P$  by  $P + 1$  diagonal matrix of the singular values of  $Y_C(\alpha)$ , denoted  $\sigma_i$  for  $i = 1 \dots P + 1$ , and arranged so that  $\sigma_1 \geq \sigma_2 \dots \geq \sigma_{P+1} > 0$ . Thus, in terms of the unitary matrix  $V(\alpha_i)$

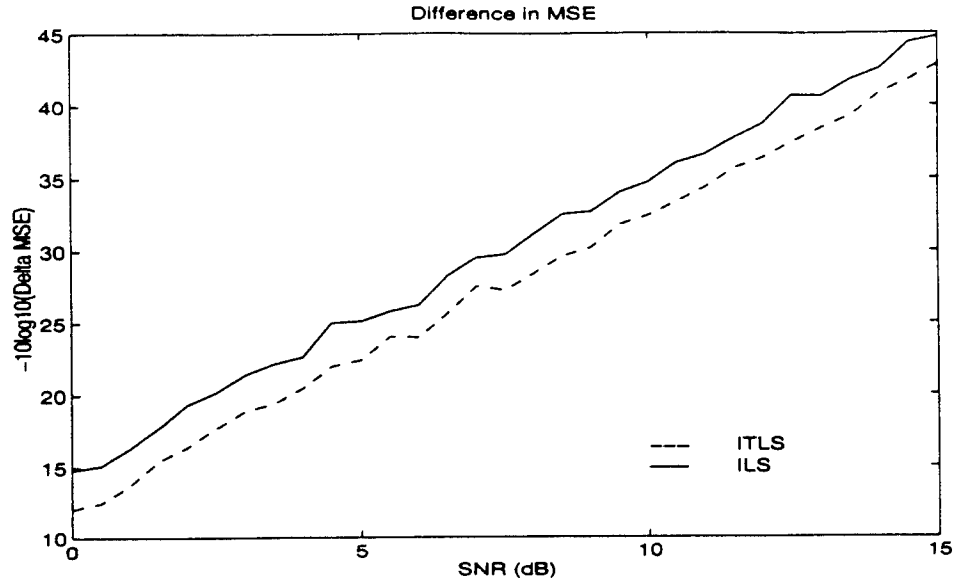
$$Y_C^T(\alpha^i)Y_C(\alpha^i) = V(\alpha^i)\Sigma^T(\alpha^i)\Sigma(\alpha^i)V^T(\alpha^i) = V(\alpha^i)\Lambda(\alpha^i)V^T(\alpha^i) \quad (3.142)$$

Here,  $\Lambda(\alpha^i)$  is the diagonal matrix of the eigenvalues associated with  $Y_C^T(\alpha^i)Y_C(\alpha^i)$ . Since  $\lambda_i = \sigma_i^2$ , the eigenvector associated with the smallest eigenvalue is  $v_{P+1}(\alpha^i)$ , the right-most singular vector obtained directly from the SVD of  $Y_C(\alpha^i)$ . That is

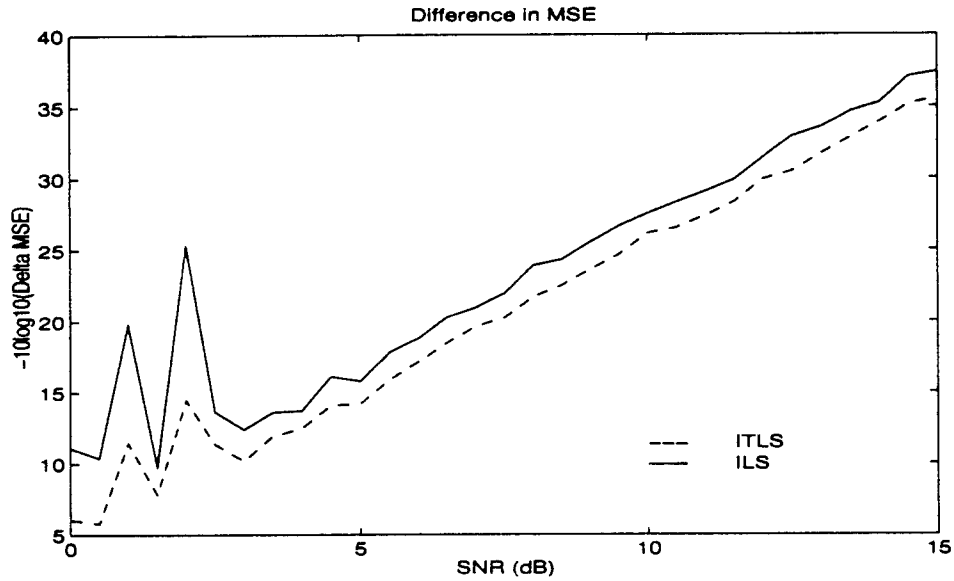
$$\alpha^{i+1} = v_{P+1}(\alpha^i) \quad (3.143)$$

This solution is exactly the same as the unnormalized fixed point solution defined by the ITLS fixed point mapping,  $\mathcal{E}_{ITLS}(\alpha^i)$ , from Equation 3.100. This derivation is an original contribution of this dissertation for it equates the IQML algorithm, previously developed as a minimization algorithm, to an ITLS fixed point solution for  $\alpha$ . What has been interpreted in the past as an approximate minimization is actually a fixed point determination for the LP coefficients.

**3.7.4 Algorithm Comparison.** Figure 3.14 shows the average difference in the squared error, at convergence, between the IEGD and ILS algorithms, and the IEGD and ITLS algorithms, for the sinusoids tested. This figure shows the ILS estimator produces, on average, a smaller squared error of the LP objective function than that produced by the ITLS algorithm. Hence, the ILS estimator will provide LP estimates closer to the ML estimates than those produced by the ITLS estimator. Figure 3.15 shows the frequency estimation accuracy of each of these algorithms for both one and two filtered sinusoids. As expected, for this data, the IEGD algorithm provides the most accurate frequency estimates. However, these estimates are only slightly better than those provided by the ILS and ITLS fixed point estimates, especially at large SNRs. This slight improvement must be weighed against the additional complexity of calculating the Hessian matrix for the IEGD algorithm at each iteration.

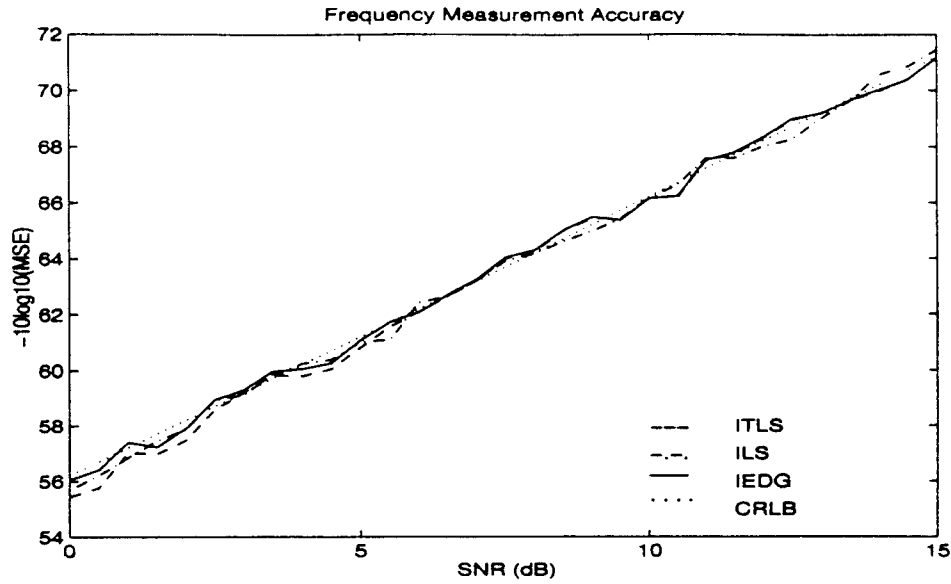


(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ .

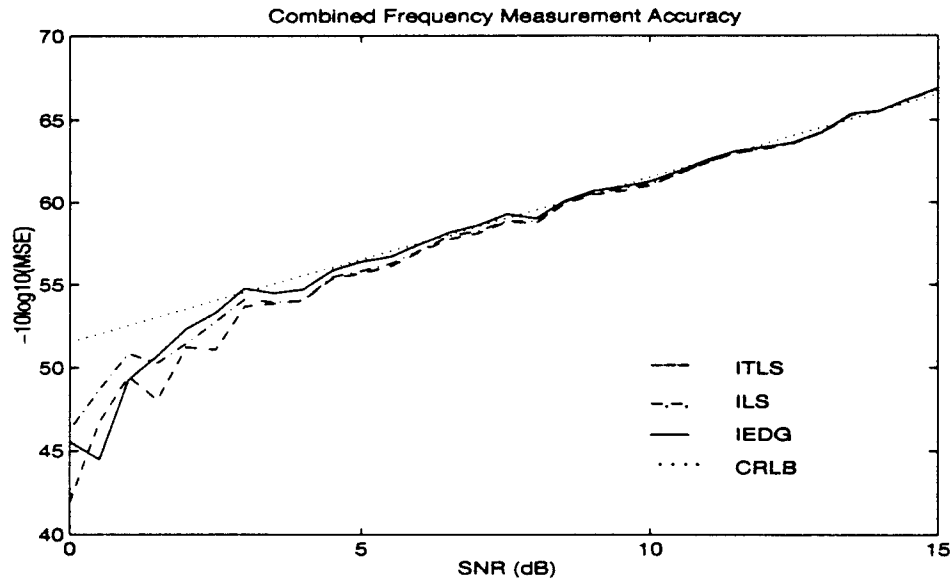


(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3], [b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .

**Figure 3.14 LP Estimator Minimum Error Comparison:**  $J(\alpha)$  evaluated at convergence for each algorithm. MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. Ten iterations allowed for convergence of IEGD, ILS and ITLS algorithms. Minimum error found from IEGD algorithm initialized via  $\alpha$  at ILS convergence. Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ . SNR calculated as  $SNR = -10\log_{10}(2\sigma^2)$ .  $MSE = \frac{1}{500} \sum_{i=1}^{500} [J(\alpha_{IEGD}) - J(\hat{\alpha}_i)]^2$ .



(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3], [b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .

Figure 3.15 **LP Estimator Frequency Estimation Comparison:** MSE calculated at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. Ten iterations allowed for convergence of each algorithm. IEDG algorithm initialized via  $\alpha$  at ILS convergence. Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$  Length:  $N = 32$ . SNR calculated as  $SNR = -10\log_{10}(2\sigma^2)$  and  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} [\hat{f}_p(i) - f_p]^2$ .

*3.7.5 Section Summary.* This section developed an exact ML estimator, termed the IEGD algorithm, for the LP coefficients based on the LP general linear model and is an original contribution of this research. The estimator derived provides ML estimates of the LP coefficients and, consequently, the frequencies of filtered sinusoids, by employing an iterative gradient descent algorithm to minimize the squared error dictated by the LP general linear model. In addition, this section derived the equivalence between the IGLS and IQML algorithms and the ILS and ITLS fixed point estimators respectively. These derivations show the IGLS and IQML algorithms are fixed point estimators and do not minimize the LP objective function in general. Simulations show the IEGD algorithm can provide ML frequency estimates by minimizing the LP objective function, though the performance increase over the less complicated ILS/IGLS and ITLS/IQML algorithms is negligible at high signal to noise ratios. In addition, for the simulations completed the ILS/IGLS algorithm actually produced, on average, more accurate frequency estimates than the more complicated ITLS/IQML algorithm. The IQML algorithm requires an eigenvalue decomposition of a  $P + 1$  by  $P + 1$  square matrix at each iteration whereas the ILS algorithm only requires the inversion of a  $P$  by  $P$  square matrix. This result is surprising since the IQML algorithm has been widely accepted as the premier method of estimating the frequencies of complex sinusoids in noise. Thus, if the small degradation in performance over the IEGD can be tolerated, the ILS/IGLS estimator should be preferred over the TLS/IQML estimator for providing minimum variance unbiased point estimates of the LP coefficients and, consequently, the sinusoidal frequencies. The next section develops a new method for determining the estimation accuracy for each point estimate of the LP coefficients and frequencies based solely on a single realization of the measurements,  $y$ .

### *3.8 Bounding the Estimation Error*

Previous portions of this dissertation have developed methods of obtaining unbiased point estimates of the LP coefficients and, consequently, the sinusoidal frequencies present in a set of  $M$  observations. However, knowledge that an estimate is unbiased is not adequate to fully characterize the estimates; the estimation error must also be quantified. For sinusoids in noise, the accuracy of each estimate is a function of several parameters includ-

ing the SNR, the number of measurements and the relative amplitudes, frequencies and phases of the sinusoids (39:414,415). This section introduces a new method for bounding the estimation error of the LP coefficient estimates and consequently, the frequency estimates, based strictly on one realization of the measurement vector. Hence, the accuracy estimate is data driven. This derivation is an original contribution of this research and can be used to establish confidence intervals for both the LP coefficients and the frequency estimates. In addition, knowledge of the accuracy of the estimate is critical for data fusion algorithms and pattern classification systems.

*3.8.1 Theory.* Let  $\theta$  represent the true values of a set of  $P$  parameters and let  $\hat{\theta}$  be an associated estimate. The estimation error of any one of the parameters, denoted  $\epsilon_p$ , will be defined as  $\theta_p$  (61:345).

$$\epsilon_p = \hat{\theta}_p - \theta_p \quad (3.144)$$

As a function of a random variable, this error is also a random variable and can best be analyzed in probabilistic terms. That is, if  $t$  is an arbitrary constant, the probability the absolute error is less than  $t$  is found as

$$P(|\epsilon_p| < t) = P(-t \leq \epsilon_p \leq t) \quad (3.145)$$

Calculation of this probability depends on the PDF of  $\epsilon_p$  which is usually unknown. Provided the estimate is unbiased, however, a loose bound on this probability of error can be found using only the variance of the estimate by employing Chebyshev's Inequality (27:58-59)

$$P(|\epsilon_p| > t) \leq \frac{V\{\hat{\theta}_p\}}{t^2} \quad (3.146)$$

Letting  $t$  be an integer multiple of the standard deviation,  $t = k\sigma_{\hat{\theta}_p}$ , yields

$$P(|\epsilon_p| \geq k\sigma_{\hat{\theta}_p}) \leq \frac{1}{k^2} \quad (3.147)$$

Thus, if the variance of an estimate for  $\theta_p$  can be calculated, a loose bound on the estimation error can be achieved. Alternatively, confidence intervals can be formed as

$$P(\hat{\theta}_p - k\sigma_{\hat{\theta}_p} < \theta_p < \hat{\theta}_p + k\sigma_{\hat{\theta}_p}) = 1 - \beta \quad (3.148)$$

Here the quantity  $(1 - \beta)$  is called the confidence coefficient and indicates the relative frequency the interval defined by the estimate and the variance of the estimate will include the true value of  $\theta_p$  (28:151-158). Again, this calculation depends on the exact form of the PDF of the error, although a bound can be obtained from Equation 3.147.

Thus, if the variance of an unbiased estimate is known or can be calculated from a single measurement vector, the accuracy of the estimate can be gauged by employing Equation 3.147 or a confidence interval can be established via Equation 3.148. A method to estimate the variance of a point estimate of the LP coefficients, using only the measurements and a knowledge of the variance of the noise, is derived in the next section.

**3.8.2 LP Coefficients.** To estimate the variance of an estimate of the LP coefficients, consider the effect of the system model on the measurements. From the LP general linear model of Equation 3.49, since  $y[m] = s[m] + w[m]$  then

$$G(\mathbf{a})Y\mathbf{a} = \mathbf{e}(\mathbf{a}) = G(\mathbf{a})S\mathbf{a} + G(\mathbf{a})A^T\mathbf{w} \quad (3.149)$$

With  $a[0] = 1$ , this form of the LP general linear model can be written as

$$[\mathbf{e}(\mathbf{a}) - G(\mathbf{a})\mathbf{s}_o] = G(\mathbf{a})S_o\mathbf{a}_o + G(\mathbf{a})A^T\mathbf{w} \quad (3.150)$$

where  $\mathbf{s}_o$  is the first column of  $S$  and  $S_o$  is an  $M - 2P$  by  $P$  matrix of the remaining columns. The ML estimate for  $\mathbf{a}_o$  is unique and found as (40:186),(60)

$$\hat{\mathbf{a}}_o(ML) = [S_o^T G^T(\mathbf{a}) G(\mathbf{a}) S_o]^{-1} \{S_o^T G^T(\mathbf{a}) [\mathbf{e}(\mathbf{a}) - G(\mathbf{a})\mathbf{s}_o]\} \quad (3.151)$$

In addition, the statistical properties of this estimate are found as (40:186)

$$E\{\hat{\mathbf{a}}_o(ML)\} = \mathbf{a}_o \quad (3.152)$$

$$K_{\hat{\mathbf{a}}_o} = [S_o^T G^T(\mathbf{a}) G(\mathbf{a}) S_o]^{-1} \quad (3.153)$$

This gives the covariance matrix for the ML estimate of the LP coefficients in terms of the deterministic model matrix  $S_o$ . Unfortunately, the elements of this matrix are derived from the deterministic signal,  $s[m]$ , which is unavailable. To estimate the covariance matrix from the observations, consider the expansion of the observation matrix defined by  $G(\mathbf{a})Y_o$ . Letting  $L = M - 2P$ , then

$$G(\mathbf{a})Y_o = \begin{bmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,L} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,L} \\ \vdots & \vdots & \cdots & \vdots \\ g_{L,1} & g_{L,2} & \cdots & g_{L,L} \end{bmatrix} \begin{bmatrix} y[M-2] & y[M-3] & \cdots & y[L-1] \\ y[M-3] & y[M-4] & \cdots & y[L-2] \\ \vdots & \vdots & \cdots & \vdots \\ y[2P-1] & y[2P-2] & \cdots & y[0] \end{bmatrix} \quad (3.154)$$

Thus, each element of this matrix product can be viewed as the output of a finite impulse response filter of length  $L$ . That is, for  $i = 1 \dots L$  and  $j = 1 \dots 2P$ ,

$$[G(\mathbf{a})Y_o]_{i,j} = \sum_{m=0}^{L-1} g_{i,m+1} y[M-m-1-j] \quad (3.155)$$

Now define the vector  $\mathbf{y}_1 = [y[M-2], \dots, y[0]]^T$  and the matrix  $G_i$  as the  $M-1$  by  $2P$  filter matrix

$$G_i^T = \begin{bmatrix} g_{i,1} & \cdots & g_{i,L} & 0 & 0 & \cdots & 0 \\ 0 & g_{i,1} & \cdots & g_{i,L} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & \vdots & 0 \\ 0 & 0 & \cdots & 0 & g_{i,1} & \cdots & g_{i,L} \end{bmatrix} \quad (3.156)$$

Then, the  $i^{th}$  row of  $G(\mathbf{a})Y_o$  can be written as  $\mathbf{y}_1^T G_i$ . Thus

$$Y_o^T G^T(\mathbf{a}) = [G_1^T \mathbf{y}_1 \dots G_L^T \mathbf{y}_1] \quad (3.157)$$

As a result, the matrix product  $[Y_o^T G^T(\mathbf{a}) G(\mathbf{a}) Y_o]$  can be written as

$$Y_o^T G^T(\mathbf{a}) G(\mathbf{a}) Y_o = \sum_{m=1}^L G_m^T \mathbf{y}_1 \mathbf{y}_1^T G_m \quad (3.158)$$

Taking the expected value of this matrix yields

$$E\{Y_o^T G^T(\mathbf{a}) G(\mathbf{a}) Y_o\} = \sum_{m=1}^L G_m^T E\{\mathbf{y}_1 \mathbf{y}_1^T\} G_m \quad (3.159)$$

Now, with  $\mathbf{s}_1$  and  $\mathbf{w}_1$  defined in a manner similar to  $\mathbf{y}_1$ , then  $\mathbf{y}_1 = \mathbf{s}_1 + \mathbf{w}_1$ , so that

$$E\{\mathbf{y}_1 \mathbf{y}_1^T\} = \mathbf{s}_1 \mathbf{s}_1^T + E\{\mathbf{w}_1 \mathbf{w}_1^T\} \quad (3.160)$$

Substitution into Equation 3.159 produces

$$E\{Y_o^T G^T(\mathbf{a}) G(\mathbf{a}) Y_o\} = \sum_{m=1}^L G_m^T \mathbf{s}_1 \mathbf{s}_1^T G_m + \sum_{m=1}^L G_m^T E\{\mathbf{w}_1 \mathbf{w}_1^T\} G_m \quad (3.161)$$

Since  $\mathbf{w}_1$  is a zero-mean, normally distributed random vector with covariance matrix  $K_w = \sigma^2 H^T H$ , then

$$E\{Y_o^T G^T(\mathbf{a}) G(\mathbf{a}) Y_o\} = S_o^T G^T(\mathbf{a}) G(\mathbf{a}) S_o + \sigma^2 \sum_{m=1}^L G_m^T H^T H G_m \quad (3.162)$$

For any ML estimate of  $\mathbf{a}$ , an estimate of the covariance matrix can be obtained as

$$\hat{K}_{\hat{\mathbf{a}}} = \left[ Y_o^T G^T(\mathbf{a}) G(\mathbf{a}) Y_o - \sigma^2 \sum_{m=1}^{M-2P} G_m^T H^T H G_m \right]^{-1} \quad (3.163)$$

Finally, employing the symmetry constraints,  $\mathbf{a} = B\boldsymbol{\alpha}$  with  $\alpha[0] = 1$ , the final form of the estimate for the covariance matrix of the constrained coefficients in  $\boldsymbol{\alpha}_o$  becomes

$$\hat{K}_{\hat{\boldsymbol{\alpha}}_o} = \left[ Y_c^T G^T(\boldsymbol{\alpha}_o) G(\boldsymbol{\alpha}_o) Y_c - \sigma^2 B^T \sum_{m=1}^{M-2P} G_m^T H^T H G_m B \right]^{-1} \quad (3.164)$$



where  $Y_c = Y_o B$ . For large signal to noise ratios, assuming  $\hat{\alpha}_o$  is relatively close to the true value of  $\alpha_o$ , an estimate of the LP coefficient covariance matrix,  $\hat{K}_{\hat{\alpha}_o}$ , becomes

$$\hat{K}_{\hat{\alpha}_o} = [Y_c^T G^T(\hat{\alpha}_o) G(\hat{\alpha}_o) Y_c]^{-1} \quad (3.165)$$

Thus, using only the measurements and exploiting a knowledge of the noise variance, the variance,  $V\{*\}$ , of each LP coefficient can be estimated as

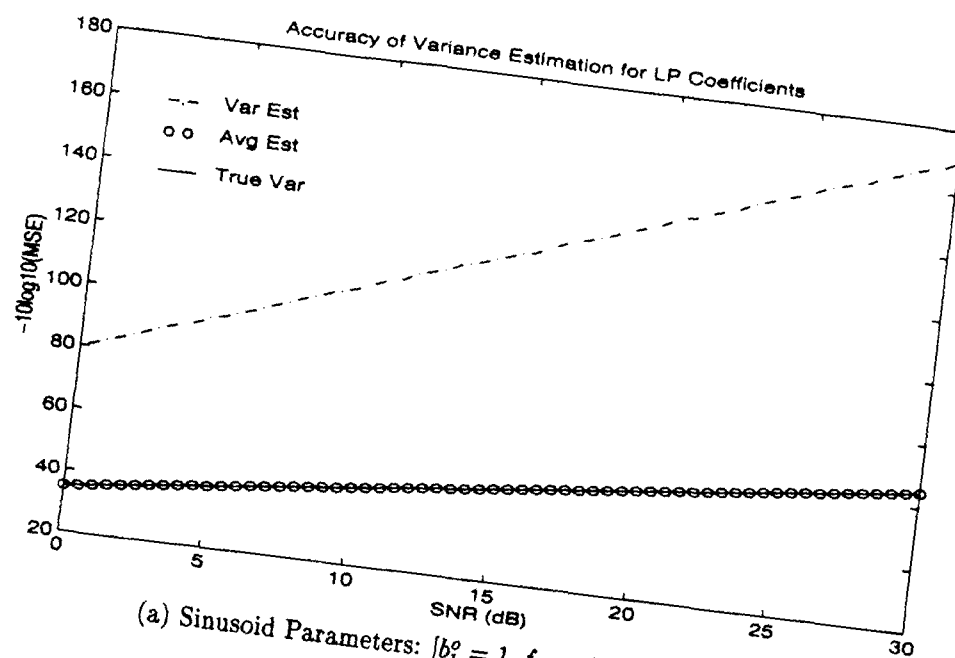
$$\hat{V}\{\alpha[p]\} = [\hat{K}_{\hat{\alpha}_o}]_{p,p} \quad (3.166)$$

Substitution of this estimate into Equation 3.147 yields a bound on the estimation error. Alternatively, substitution of this estimate into Equation 3.148 provides a confidence interval.

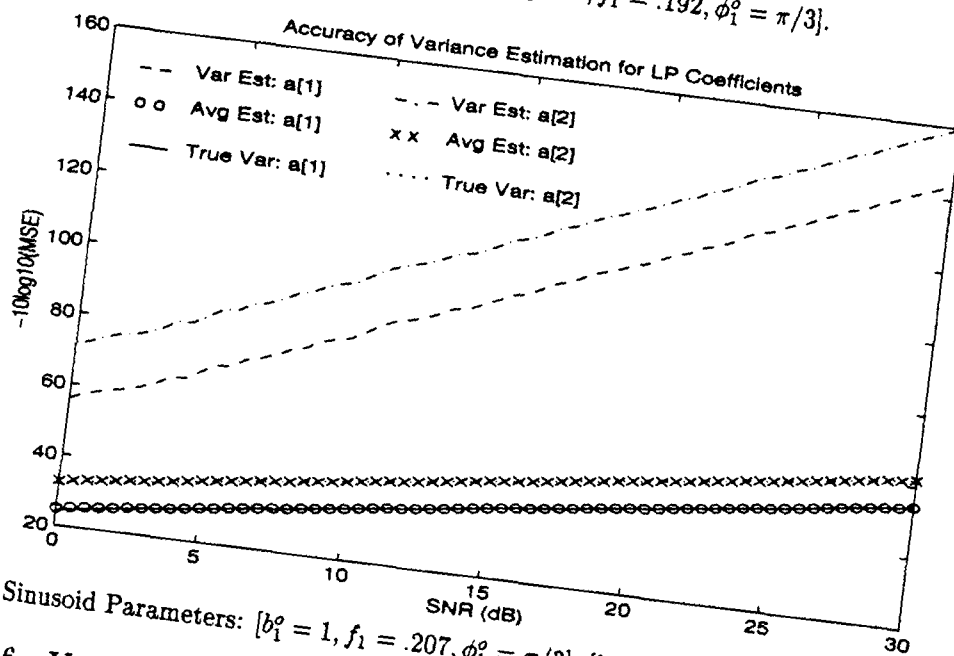
Figure 3.16 shows the accuracy of the variance estimates of the LP coefficients obtained from the ILS estimator for one and two filtered sinusoids. For both cases, the average value of the variance estimates coincides with the true variance of the LP estimates as given by the CRLB. In addition, the variance of these variance estimates is extremely small. This indicates each estimate of the variance will be relatively close to the true value of the variance so that Equation 3.147 and Equation 3.148 can be used to accurately characterize the accuracy of a point estimate of the LP coefficients. This information can also be used to determine the accuracy of the associated frequency estimates.

**3.8.3 Sinusoidal Frequencies.** Previously, a method was derived to estimate the variance of the LP coefficients from a single measurement vector and knowledge of the noise variance. To map this estimate to the frequency space, a transformation must be employed. In general, let  $\hat{\alpha}$  be an unbiased estimate of  $\alpha$  and let  $K_{\hat{\alpha}}$  be the covariance matrix of the estimate. In addition, let  $f_p = g_p(\alpha)$  be a scalar function of  $\alpha$ . The variance of  $f_p$  can be estimated as (19:183),(20).

$$V\{f_p\} \approx \left[ \frac{\partial g_p(\alpha)}{\partial \alpha} \right]^T K_{\hat{\alpha}} \left[ \frac{\partial g_p(\alpha)}{\partial \alpha} \right] \bigg|_{\alpha=\hat{\alpha}} \quad (3.167)$$



(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3], [b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .

**Figure 3.16 Variance Estimation of LP Coefficients:** Variance estimated at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. LP coefficients estimated by ILS algorithm. Ten iterations allowed convergence. True variance found from CRLB. Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ . SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ .  

$$\text{Var Est} = \frac{1}{500} \sum_{i=1}^{500} [\hat{V}\{a[p]\}_i - V\{a[p]\}_{TRUE}]^2$$

$$\text{Avg Est} = \frac{1}{500} \sum_{i=1}^{500} \hat{V}\{a[p]\}_i$$

Thus, for a scalar function,  $g_p(\alpha)$ , an estimate of the variance can be obtained provided the first derivatives exist and  $\hat{\alpha}$  is not too far from  $\alpha$ .

In general, there is no closed form function which relates the coefficients  $\{a[p]\}$  directly to the frequencies  $\{f_p\}$ . However, for  $P = 1$  and  $P = 2$ , and to a certain extent for  $P = 3$  and  $P = 4$ , closed form functions relating the frequencies to the coefficients can be derived which allow Equation 3.167 to be employed to bound the estimation error for the frequencies.

For one sinusoid, from Equation 3.57, From Equations the function relating  $f_1$  to the coefficient  $a[1]$  is given as

$$f_1 = \frac{1}{2\pi} \cos^{-1} \left[ \frac{-a[1]}{2} \right] = g_1(a[1]) \quad (3.168)$$

On the interval  $0 \leq f_1 < .5$ , then  $-2 < a[1] < 2$  and the variance of the frequency estimate becomes (48:462)

$$V\{f_1\} \approx \frac{1}{4\pi^2(4 - a[1]^2)} V\{a[1]\} \Big|_{a[1]=\hat{a}[1]} \quad (3.169)$$

For two sinusoids, employing Equation 3.58 yields the relationships (53)

$$f_p = g_p(\alpha_o) = \frac{1}{2\pi} \cos^{-1}(u_p) \quad (3.170)$$

where the quantity  $u_p$  is defined as

$$u_p = \frac{1}{4} \left[ -a[1] + (-1)^{p+1} \sqrt{a[1]^2 - 4a[2] + 8} \right] \quad (3.171)$$

for  $p = 1, 2$ , provided  $a[1]^2 - 4a[2] + 8 \geq 0$ . Substitution for  $a[p]$  yields

$$[\cos(2\pi f_1) - \cos(2\pi f_2)]^2 \geq 0 \quad (3.172)$$

Since this inequality holds for all  $f$ , the mapping between  $f$  and  $\alpha_o$  exists. Taking the derivative with respect to each parameter yields

$$\frac{\partial g_p(\alpha_o)}{\partial \alpha[1]} = \frac{1}{8\pi\sqrt{1-u_p^2}} \left[ -1 - \frac{(-1)^p a[1]}{\sqrt{a[1]^2 - 4a[2] + 8}} \right] \quad (3.173)$$

$$\frac{\partial g_p(\alpha_o)}{\partial \alpha[2]} = \frac{1}{8\pi\sqrt{1-u_p^2}} \left[ 2 \frac{(-1)^p}{\sqrt{a[1]^2 - 4a[2] + 8}} \right] \quad (3.174)$$

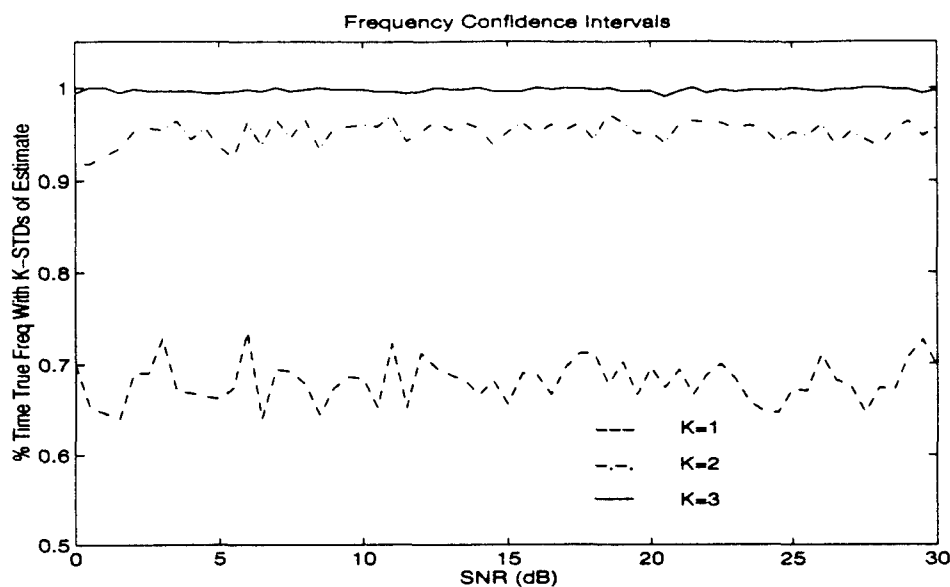
Thus, the variance of the frequency estimates becomes

$$\hat{V}\{\hat{f}_p\} = \left[ \frac{g_p(\hat{\alpha}_o)}{\partial \alpha_o} \right]^T \hat{K}_{\alpha_o} \left[ \frac{g_p(\hat{\alpha}_o)}{\partial \alpha_o} \right] \quad (3.175)$$

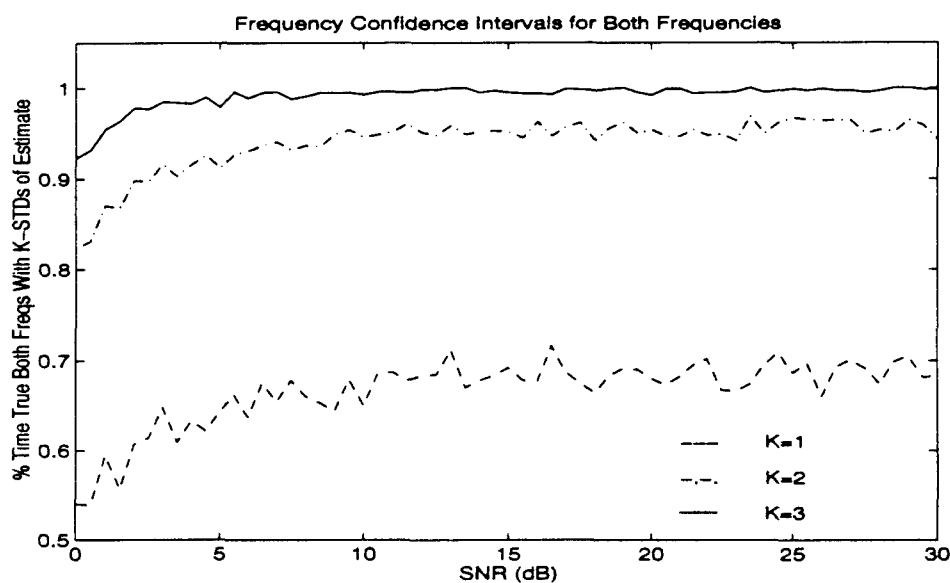
These variance estimates can now be used to bound the estimation error or develop confidence intervals for each point estimate of the frequencies.

Figure 3.17 shows the accuracy of the frequency confidence intervals, obtained from Equation 3.148, for both one and two filtered sinusoids. Here, the ILS algorithm was used to determine the LP coefficients while the variance estimates for the LP coefficients were obtained from Equation 3.165. The variance estimates for the frequencies were calculated from Equation 3.175. For this data, Figure 3.17 shows the percentage of time the confidence intervals, developed from the point estimates, enclosed the true value of the frequency for one, two and three standard deviations of the variance estimates. In each case, above the SNR threshold, the confidence intervals established provide an accurate method of gauging the accuracy of the point estimate of the frequency.

**3.8.4 Section Summary.** This section derived a method for bounding the estimation error of the LP coefficient estimates and consequently, the frequency estimates, based strictly on one realization of the measurement vector, and is an original contribution of this dissertation. Specifically, from the LP general linear model, this section derived a method for estimating the variance of the LP coefficients based solely upon knowledge of the noise variance and the set of measurements. This unbiased estimate of the LP variance was then transformed into an estimate of the frequency variance for one and two sinusoids. Simulations indicate these estimates of the measurement variance can be used to gauge



(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .192, \phi_1^o = \pi/3]$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .207, \phi_1^o = \pi/3], [b_2^o = 1, f_2 = .227, \phi_2^o = 4\pi/3]$ .

**Figure 3.17 Frequency Confidence Intervals:** Confidence intervals established at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. LP coefficients estimated by ILS algorithm. Ten iterations allowed convergence. Filter Parameters: Center Frequency:  $f_c = .21$ ; Bandwidth:  $f_B = .2$ ; Length:  $N = 32$ . Confidence interval:  $|\hat{f}_p(i) - f_p| < k\hat{\sigma}_f$  for  $k = 1, 2, 3$ . SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ .

the accuracy of the point estimates of both the LP coefficients and the frequencies. This gauge can then be employed to develop confidence intervals to aid any decision making process based on a single set of measurements.

### *3.9 Summary*

This chapter derived estimators for the amplitude, phase and frequency parameters of filtered sinusoids in noise. After deriving a mathematical model describing the effects of a filter on sinusoids and noise, ML estimators for the sinusoidal parameters were developed based on this model. During this development, it was shown that ML estimates of the amplitude and phase parameters require ML estimates of the frequencies in addition to the use of the correct system model. An ML estimator for the frequencies was then derived based on an iterative gradient search employing the ML frequency objective function. Simulations indicate this technique can produce minimum variance unbiased frequency estimates provided the effects of the filter are correctly incorporated into the system model.

Since this ML frequency estimator can require many iterations to achieve convergence, an alternative method for estimating the frequencies was developed based on linear predictive modeling. Specifically, an equivalent general linear model, parameterized by a set of LP coefficients was derived for estimating the frequencies of the sinusoids in noise. The LP general linear model, derived in this chapter, properly accounts for the effects of the LP coefficients and system filter on the noise and is an original contribution of this research. Estimation of the sinusoidal frequencies was recast as the estimation of the LP coefficients using the general linear model.

The ILS and ITLS estimators were then derived, based on fixed-point theory, to estimate the LP coefficients of the general linear model from the measurements. The development of these estimators represents an original contribution of this research. Simulations indicate each of these estimators produce accurate estimates of the LP coefficients and, consequently the frequencies, above a certain SNR threshold; convergence is attained in a few iterations.

This chapter then related the ML frequency objective function for filtered sinusoids to the LP general linear model. The analysis showed that the LP coefficients which minimize the sum of the squared errors as defined by the LP general linear model also provide ML frequency estimates and is an original contribution of this research. An ML estimator for the LP coefficients, termed the IEGD algorithm, was then derived based on an iterative minimization of the LP general linear model squared error and is an original contribution of this research. Simulations indicate this algorithm will provide the LP coefficients producing the minimum squared error and consequently, ML estimates of the LP coefficients and frequencies.

Since the IEGD algorithm can be computationally difficult to implement, the IGLS and IQML approximate minimization algorithms were examined. Analysis revealed that the IGLS algorithm is equivalent to the ILS fixed point estimator of the LP coefficients while the IQML algorithm is equivalent to the ITLS fixed point estimator. This recasting of the IGLS and IQML algorithms as fixed point estimators is an original contribution of this research. Simulations comparing the performance of the ILS and TLS algorithms with the IEGD algorithm indicate the IEGD algorithm produces frequency estimates which are only slightly better than those obtained via ILS and ITLS. In addition, the simulations indicated the ILS algorithm produces frequency estimates at least as accurate as those produced by the ITLS algorithm. All other factors being equal, the ILS algorithm is thus preferred over the ITLS algorithm since the ITLS algorithm requires an iterative SVD of the observation matrix.

Finally, this chapter derived a method of bounding the estimation error of the point estimates of the LP coefficients and the frequencies based only on the measurements and a knowledge of the noise variance. This derivation employed the LP general linear model to obtain unbiased estimates of the covariance matrix of the LP coefficients and is an original contribution of the research. Simulations indicate the estimates of the variance are good enough to allow confidence intervals to be established for the frequencies.

## IV. Application: Digital Electronic Warfare Receiver

### 4.1 Introduction

This chapter applies the ILS fixed point estimation technique to the digital EW receiver being built by the Air Force Research Laboratory/Wright Laboratory (AFRL/WL). After discussing the general receiver architecture, models of the real and complex data output from the receiver will be constructed. A method to efficiently implement the receiver as a nonmaximally decimated, Uniform Discrete Fourier Transform (UDFT) polyphase filter bank is then derived. After showing there is no loss in frequency estimation accuracy attributable to the use of the complex data model as opposed to the real data model, a new form of the LP general linear model is derived for complex sinusoids. Identification of the LP coefficients for this model is then shown to be related to ML frequency estimation and an ILS fixed point estimator is derived for estimating the LP coefficients. After applying the ILS estimator to the receiver output for both one and two sinusoids, the chapter concludes by showing the PDF of the complex filtered data can be approximated as a complex multivariate Gaussian PDF. This approximation, in turn, allows a simplified frequency estimator to be derived for a single complex sinusoid.

### 4.2 Digital Receiver Architecture

As shown in Figure 4.1, the basic structure of a digital channelized receiver can be divided into the antialiasing filter, Analog-to-Digital Converter (ADC), filter bank, demodulators, decimators and parameter encoders.

In general, the input signal can be modeled as the sum of  $P_o$  analog pulsed sinusoids in zero-mean, normally distributed, independent noise,  $\eta(t)$

$$x(t) = \nu(t) + \eta(t) = \sum_{p=1}^{P_o} b_p^o \cos(2\pi F_p t + \phi_p^o) + \eta(t) \quad (4.1)$$

Here,  $F_p$  is the analog frequency,  $\phi_p^o$  is the phase, and  $b_p^o$  is the amplitude of the  $p^{th}$  sinusoid. This analog signal is passed through a 34<sup>th</sup> order Chebyshev bandpass antialiasing filter with critical frequencies as shown in Table 4.1 (16). This filter reduces the effects of



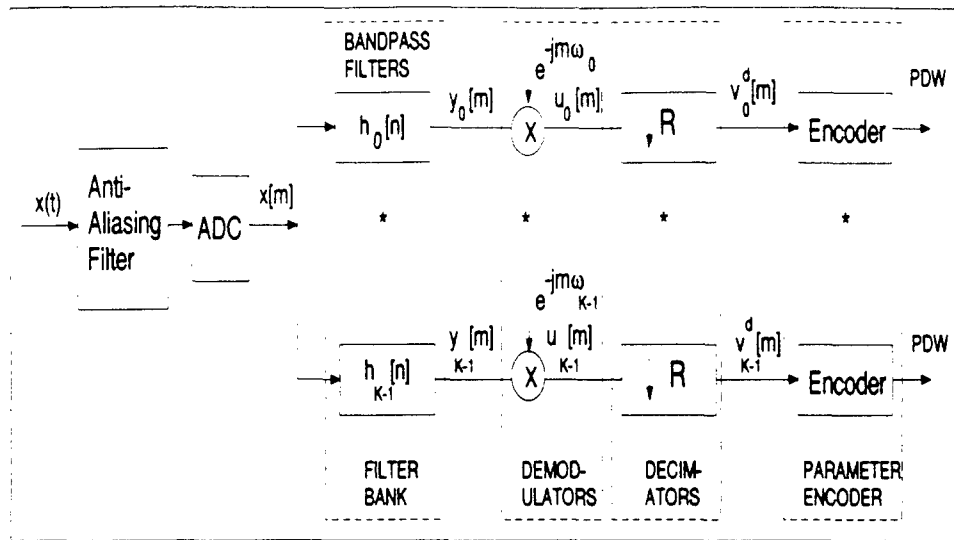


Figure 4.1 **EW Digital Receiver Architecture:**  $K = 32$  Bandpass Filters;  $R = 16$  Decimation Rate; 8-bit ADC with sample rate  $f_s = 3\text{GHz}$ ; Chebyshev Antialiasing Filter with 1.5 GHz Bandwidth.

out-of-band interfering signals and noise while passing signals within a specific frequency range for digitization (56:2). For bandpass sampling of real signals, the 1.5 GHz operational bandwidth, (BW), of this filter is related to the sample rate,  $f_s$ , of the ADC as  $2BW \leq f_s \leq 4BW$  (84:125).

An analog signal passing through this filter is sampled by the eight bit ADC at a sample rate of  $f_s = 3 \times 10^9$  samples per second. The resulting digital signal,  $x[m]$ , becomes (64:10)

$$x[m] = \sum_{p=1}^{P_o} b_p^o \cos(2\pi f_p m + \phi_p^o) + \eta[m] \quad (4.2)$$

where the digital frequency,  $f_p$ , is related to the analog frequency,  $F_p$ , via the transformation  $f_p = F_p/f_s$ . Thus, the analog frequency range  $1.5\text{GHz} \leq F_p \leq 3\text{GHz}$  is inversely mapped to the digital range  $0 \leq f_p \leq .5$ . Finally, the noise,  $\eta[m]$ , is zero mean and normally distributed with a covariance matrix determined by the characteristics of the antialiasing filter.

In order to provide for broadband instantaneous frequency coverage and allow for simultaneous signal detection and estimation, the operational bandwidth is divided into  $K = 32$  specific frequency bands by employing a set of bandpass filters. The coefficients

Table 4.1 Filter Critical Frequencies

Filter Type	Order	Freq	Gain
Anti-aliasing	34	1.4953(GHz) and 3.047(GHz)	-60dB
		1.5047(GHz) and 2.953(GHz)	-3dB
Prototype	192	1/32 rads/sec	-60dB
		1/64 rads/sec	-3dB
$k^{th}$ Bandpass	192	$k/32 \pm 1/32$	-60dB
		$k/32 \pm 1/64$	-3dB

for the  $k^{th}$  bandpass filter,  $\tilde{h}_k[n]$ , are complex and derived from a real, symmetric, lowpass filter,  $h_0[n]$ , by

$$\tilde{h}_k[n] = h_0[n]e^{j2\pi kn/K} \quad (4.3)$$

This prototype filter,  $h_0[n]$ , is an  $N = 192$  tap, real, symmetric filter designed using the McClellan-Parks algorithm to meet the frequency characteristics given in Table 4.1. Since  $H_k(e^{j2\pi f}) = H_0(e^{j[2\pi f - 2\pi f_k]})$ , each channel of the filter bank will be a bandpass filter with  $|H_k(e^{j2\pi f})| = 0$  for  $|f - f_k| > 1/32$ . For a real signal,  $x[m]$ , the output of the  $k^{th}$  filter,  $\tilde{y}_k[m]$ , will be complex and given as the convolution sum

$$\tilde{y}_k[m] = \sum_{n=0}^{N-1} h_0[n]x[m-n]e^{j2\pi kn/K} = \sum_{l=0}^{K-1} \sum_{n=0}^{N_o-1} h_0[l+nK]x[m-l-nK]e^{j2\pi kl/K} \quad (4.4)$$

where  $N = KN_o$  for integer  $N_o$ . As shown in Appendix F,  $\tilde{y}_k[m]$  can be interpreted as the Short Time Fourier Transform (STFT) of  $x[m]$  as seen by a causal finite window  $q[n] = h_0[N-1-n]$  and evaluated at  $\omega = 2\pi k/K$ . Conversely, a real output,  $y_k[m]$ , can be obtained by adding the outputs from channels  $k$  and  $K-1-k$ . In either case, since the spectrum will be band-limited, the output of each filter can be translated to standard frequency range via a set of modulators so that

$$\tilde{u}_k[m] = \tilde{y}_k[m]e^{-j\omega_1 m} \quad (4.5)$$

This demodulation allows for a standard set of signal detection and parameter extraction techniques to be developed for each channel.

For EW applications, the speed at which the sampled data can be processed is much slower than the data rate of the signals. In order to compensate, after demodulation, the outputs of each channel can be decimated by an integer factor,  $R$ , so that  $\tilde{v}_k^d[m] = \tilde{u}_k[Rm]$ . Consequently, the spectrum of  $\tilde{v}_k^d[m]$  can be written as(64:102-102)

$$V_k^d(e^{j\omega}) = \frac{1}{R} \sum_{r=0}^{R-1} U_k \left( \exp \left\{ \frac{j(\omega - 2\pi r)}{R} \right\} \right) \quad (4.6)$$

This decimation process serves to scale the frequency spectrum of  $v_k[m]$ .

After decimation, the output of each channel is passed to a parameter encoder which estimates the parameters, such as amplitude, frequency and phase, of a pulsed sinusoid within the channel and outputs a Pulse Descriptor Word (PDW) containing these estimates. These estimates, in turn, provide the data needed for a classification system to determine the type of radar emitting the pulse. Before constructing the parameter encoder, the output data must first be accurately modeled.

**4.2.1 Real Data Model.** Since all  $K = 32$  filters are derived from a single prototype and the output of each channel is modulated to a standard frequency band prior to decimation, the output of each channel will be the same as the output,  $y[m]$ , of a real filter,  $h[n]$  with a bandwidth of  $1/16$ , followed by an  $R$ -fold decimator. For the sampled signal,  $x[m]$ , using the results from Chapter III, an expression for the decimated output can be written in the form

$$y[m] = s[Rm] + w[Rm] = \sum_{p=1}^P b_p \cos(2\pi R f_p m + \phi_p) + w[Rm] \quad (4.7)$$

The decimator has effectively decorrelated the output noise by the decimation factor  $R$  and scaled each sinusoidal frequency to  $Rf_p$ . To prevent aliasing and allow for accurate frequency estimation,  $R < .5K = 16$ . In addition, since several of the sinusoids will have been greatly attenuated by the filter, the summation is now only over  $P$  where  $P \leq P_o$ . The problem now is to estimate the amplitudes,  $\{b_p\}$ , phases,  $\{\phi_p\}$ , and frequencies,  $\{f_p\}$ , of the sinusoids given the decimated measurements.

Now assume there are  $M$  samples of the decimated steady state output and define  $\mathbf{h}$  as a vector of filter coefficients by  $\mathbf{h} = [h[0], \dots, h[N-1]]^T$  and the vector of  $R$  zeros as  $\mathbf{0}_R = [0 \dots 0]^T$ . Using these definitions, the  $R(M-1) + N$  by  $M$  filter matrix,  $H$ , can be written as

$$H = \begin{bmatrix} \mathbf{h} & \mathbf{0}_R & \dots & \mathbf{0}_R \\ \mathbf{0}_R & \mathbf{h} & \dots & \mathbf{0}_R \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_R & \mathbf{0}_R & \dots & \mathbf{h} \end{bmatrix} \quad (4.8)$$

For  $m = 0 \dots M-1$ , a vector form of the decimated output can be written as

$$\mathbf{y} = H^T \mathbf{x} = H^T \boldsymbol{\nu} + H^T \boldsymbol{\eta} = \mathbf{s} + \mathbf{w} \quad (4.9)$$

The decimated noise vector,  $\mathbf{w}$ , is a zero-mean, normally distributed, correlated noise sequence with covariance  $K_{\mathbf{w}} = \sigma^2 H^T H$ . Again, using the results of Chapter III, define the deterministic signal matrix,  $\Lambda_{\mathcal{R}}$ , as  $\Lambda_{\mathcal{R}} = [\Lambda_1 | \Lambda_2]$  where the  $M$  by  $P$  matrices,  $\Lambda_1$  and  $\Lambda_2$ , are defined by

$$\begin{aligned} [\Lambda_1]_{l,p} &= \cos(2\pi f_p R[M-l+N-1]) \\ [\Lambda_2]_{l,p} &= -\sin(2\pi f_p R[M-l+N-1]) \end{aligned} \quad (4.10)$$

for  $l = 1 \dots M$  and  $p = 1 \dots P$ . With scaling vectors  $\mathbf{b}_s = [b_1 \sin(\phi_1) \dots b_P \sin(\phi_P)]^T$  and  $\mathbf{b}_c = [b_1 \cos(\phi_1) \dots b_P \cos(\phi_P)]^T$ , a vector form of the decimated output becomes

$$\mathbf{y} = \Lambda_1 \mathbf{b}_c + \Lambda_2 \mathbf{b}_s + \mathbf{w} = \Lambda_{\mathcal{R}} \mathbf{b} + \mathbf{w} \quad (4.11)$$

where  $\mathcal{R}$  denotes the real data model. Here  $\mathbf{b}^T = [\mathbf{b}_c^T; \mathbf{b}_s^T]$  and the  $2P$  columns of  $\Lambda_{\mathcal{R}}$  are linearly independent. This is equivalent to viewing  $\mathbf{y}$  as a set of observations generated from the deterministic system,  $\Lambda_{\mathcal{R}} \mathbf{b}$ , and corrupted by  $\mathbf{w}$ . As such, the PDF for this model becomes

$$p(\mathbf{y}; \boldsymbol{\theta}) = [2\pi]^{-5M} |K_{\mathbf{w}}|^{-5} \exp\left[-\frac{1}{2}(\mathbf{y} - \Lambda_{\mathcal{R}} \mathbf{b})^T K_{\mathbf{w}}^{-1} (\mathbf{y} - \Lambda_{\mathcal{R}} \mathbf{b})\right] \quad (4.12)$$

where  $\theta^T = [b^T, f^T]$ . As in Chapter III, the problem of estimating the frequencies, phases, and amplitudes has been changed to the problem of estimating the deterministic signal matrix  $\Lambda_R$  and the scaling coefficients  $b_c$  and  $b_s$ . The development of this model for real filtered and decimated data is an original contribution of this research.

**4.2.2 Complex Data Model.** Since the input signal is real, the outputs from channels  $0, \dots, \frac{K}{2} - 1$  contain the same information as the outputs from channels  $\frac{K}{2}, \dots, K - 1$ . Thus, an alternative model for the data can be constructed by forming a complex model for the data from channels  $0, \dots, K/2 - 1$ . In addition, since each bandpass filter is formed by modulating the prototype filter with a complex exponential and each filter output is demodulated to a specific frequency band prior to decimation, it is possible to build a single complex model for each channel based on a standard model. That is, with  $x[m]$  being the sum of  $P_o$  real sinusoids in noise and  $\tilde{h}[n]$  representing a complex filter band-limited to  $1/16$ , the decimated output,  $\tilde{y}[m]$ , of any given channel can be written as

$$\tilde{y}[m] = \tilde{s}[Rm] + \tilde{w}[Rm] \quad (4.13)$$

Here  $\tilde{s}[m]$  and  $\tilde{w}[m]$  represent the convolution of the real sinusoids,  $\nu[m]$ , and noise,  $\eta[m]$ , with the complex filter. For steady state conditions,

$$\tilde{s}[m] = \frac{1}{2} \sum_{p=1}^{P_o} b_{1p} e^{j\omega_p m} e^{j\phi_{1p}} + b_{2p} e^{-j\omega_p m} e^{j\phi_{2p}} \quad (4.14)$$

where  $b_{1p} = b_p^o |H(e^{j\omega_p})|$ ;  $\phi_{1p} = \Phi(\omega_p) + \phi_p$ ; and  $b_{2p} = b_p^o |H(e^{-j\omega_p})|$  while  $\phi_{2p} = \Phi(-\omega_p) - \phi_p$ . Since the filter is band-limited, assuming only  $P$  complex sinusoids are within the filter passband, the complex form of the output signal,  $\tilde{y}[m]$ , becomes

$$\tilde{y}[m] = \sum_{p=1}^P b_p e^{j2\pi R f_p m} e^{j\phi_p} + \tilde{w}[Rm] \quad (4.15)$$

The decimator has effectively decorrelated the output noise by the decimation factor  $R$  and scaled each sinusoidal frequency to  $Rf_p$ . To prevent aliasing and allow for accurate frequency estimation,  $R < K = 32$ . Employing the complex data model for the output of

each channel yields a method to increase the decimation rate,  $R$ , by a factor of 2 over the real data model without losing information and is an original contribution of this research.

For this model, the complex noise output,  $\tilde{w}[m]$ , is a random variable which must be described statistically. That is,  $\tilde{w}[Rm] = w_r[Rm] + jw_i[Rm]$  where  $w_r[m]$  and  $w_i[m]$  represent the convolution of the real noise with the real and imaginary components of the filter. Assuming there are  $M$  samples of the decimated steady state output, with the vectors of filter coefficients defined as  $\mathbf{h}_r = [h_r[0], \dots, h_r[N-1]]^T$ , and  $\mathbf{h}_i = [h_i[0], \dots, h_i[N-1]]^T$ , the  $R(M-1) + N$  by  $M$  filter matrices,  $H_r$  and  $H_i$  can be written as

$$H_r = \begin{bmatrix} \mathbf{h}_r & \mathbf{0}_R & \dots & \mathbf{0}_R \\ \mathbf{0}_R & \mathbf{h}_r & \dots & \mathbf{0}_R \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_R & \mathbf{0}_R & \dots & \mathbf{h}_r \end{bmatrix} \quad H_i = \begin{bmatrix} \mathbf{h}_i & \mathbf{0}_R & \dots & \mathbf{0}_R \\ \mathbf{0}_R & \mathbf{h}_i & \dots & \mathbf{0}_R \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_R & \mathbf{0}_R & \dots & \mathbf{h}_i \end{bmatrix} \quad (4.16)$$

Then, in vector form  $\mathbf{w}_r = H_r^T \boldsymbol{\eta}$  where  $\mathbf{w}_r = [w_r[M-1] \dots w_r[0]]^T$  and  $\mathbf{w}_i = H_i^T \boldsymbol{\eta}$  where  $\mathbf{w}_i = [w_i[M-1] \dots w_i[0]]^T$ . The real, augmented noise vector  $\mathbf{w}^T = [\mathbf{w}_r^T, \mathbf{w}_i^T]$  is a normally distributed, zero mean random vector with covariance matrix

$$K\mathbf{w} = \sigma^2 \begin{bmatrix} H_r^T H_r & H_r^T H_i \\ H_i^T H_r & H_i^T H_i \end{bmatrix} \quad (4.17)$$

In general,  $H_i^T H_i \neq H_r^T H_r$  and  $H_r^T H_i \neq -H_i^T H_r$  so that  $\tilde{\mathbf{w}}$  cannot be formally represented as a complex normally distributed random variable as defined in Appendix H.

To accurately model the complex data,  $\tilde{\mathbf{y}}$ , rewrite the vector in terms of the concatenation of the real vector,  $\mathbf{y}_r$ , and the imaginary vector,  $\mathbf{y}_i$ , as  $\mathbf{y}^T = [\mathbf{y}_r^T; \mathbf{y}_i^T]$  where  $\mathbf{y}_r = [y_r[M-1] \dots y_r[0]]^T$  and  $\mathbf{y}_i = [y_i[M-1] \dots y_i[0]]^T$ . Expanding yields

$$\begin{aligned} y_r[m] &= \sum_{p=1}^P b_p^c \cos(2\pi R f_p m) - b_p^s \sin(2\pi R f_p m) + w_r[Rm] \\ y_i[m] &= \sum_{p=1}^P b_p^c \sin(2\pi R f_p m) + b_p^s \cos(2\pi R f_p m) + w_i[Rm] \end{aligned} \quad (4.18)$$

As with the real data model,  $b_p^c = b_p \cos(\theta_p)$  and  $b_p^s = b_p \sin(\theta_p)$  are real scaling coefficients. Finally, with  $\Lambda_1$  and  $\Lambda_2$  defined be the  $M$  by  $P$  real signal matrices of Equation 4.10 then

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_r \\ \mathbf{y}_i \end{bmatrix} = \begin{bmatrix} \Lambda_1 & \Lambda_2 \\ -\Lambda_2 & \Lambda_1 \end{bmatrix} \begin{bmatrix} \mathbf{b}_c \\ \mathbf{b}_s \end{bmatrix} + \begin{bmatrix} \mathbf{w}_r \\ \mathbf{w}_i \end{bmatrix} = \Lambda_C \mathbf{b} + \mathbf{w} \quad (4.19)$$

where  $\Lambda_C$  is the deterministic signal matrix for the complex model. Here,  $\mathbf{y}$  and  $\mathbf{w} \in R^{2M}$ , while  $\Lambda_C \in R^{2M \times 2P}$ . The PDF of  $\mathbf{y}$  becomes

$$p(\mathbf{y}; \boldsymbol{\theta}) = [2\pi]^{-M} |\mathbf{K}\mathbf{w}|^{-.5} \exp\left[-\frac{1}{2}(\mathbf{y} - \Lambda_C \mathbf{b})^T \mathbf{K}\mathbf{w}^{-1}(\mathbf{y} - \Lambda_C \mathbf{b})\right] \quad (4.20)$$

This equation accurately models complex data derived from passing a real signal through a complex filter and is an original contribution of this research. As with the real data model, the problem of estimating the frequencies, phases, and amplitudes has been changed to the problem of estimating the signal matrix  $\Lambda$  and the scaling coefficients  $\mathbf{b}_c$  and  $\mathbf{b}_s$ .

**4.2.3 Receiver Implementation.** In order to exploit the advantages of the decimation process, the digital EW receiver described above is implemented as a multirate, nonmaximally decimated, polyphase, UDFT filter bank where, in general, the number of channels  $K$ , is an integer multiple of the decimation rate  $R$ . This unique implementation allows every operation after the ADC to be performed at a rate of  $1/R$  the rate suggested by Figure 4.1.

To accomplish this data rate reduction, the prototype filter,  $h_0[n]$  must decomposed into  $K$  polyphase components  $E_k(z)$  as (92)

$$H_0(z) = \sum_{l=0}^{K-1} z^{-l} E_l(z^K) \quad (4.21)$$

Here,  $E_l(z^K) = \sum_{n=0}^{N_o-1} h_o[nK + l]z^{-nK}$  and the integer  $N_o$ , defined by  $KN_o = N$ , is the length of the  $l^{th}$  polyphase filter. As shown in Figure 4.2, with the Inverse Discrete Fourier Transform (IDFT) matrix,  $W_{IDFT}$ , defined by  $[W_{IDFT}]_{k,l} = e^{j2\pi kl/K}$ , the output of the  $k^{th}$

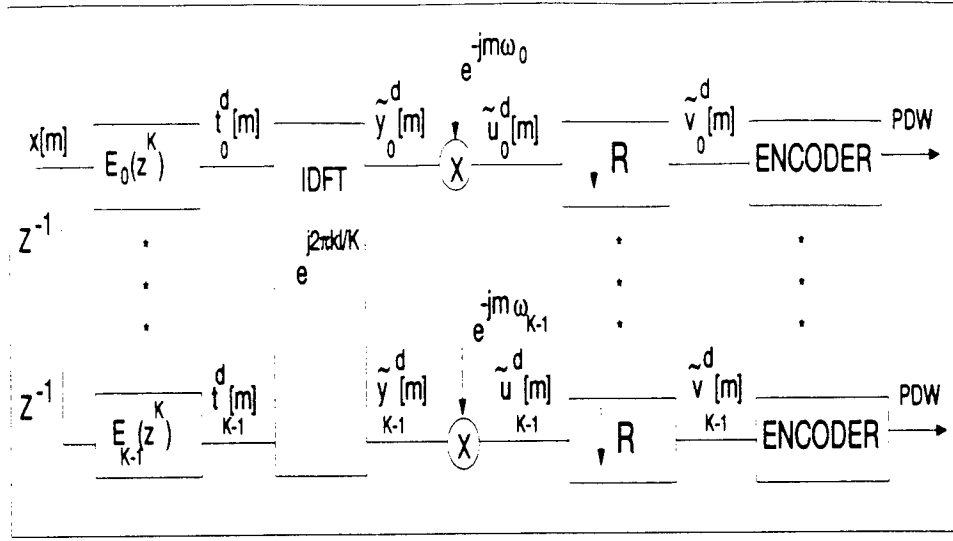


Figure 4.2 UDFT Filter Bank Polyphase Implementation

channel is given as

$$\tilde{y}_k[m] = \sum_{l=0}^{K-1} t_l[m] \exp\{j2\pi lk/K\} \quad (4.22)$$

where  $t_l[m]$  is the output of the  $l^{th}$  polyphase component. Thus, the z-transform of  $\tilde{y}_k[m]$  becomes

$$\tilde{Y}_k(z) = \sum_{l=0}^{K-1} z^{-l} \exp\{j2\pi kl/K\} E_l(z^K) X(z) \quad (4.23)$$

so that the transfer function for the  $k_{th}$  channel can be written as

$$\tilde{H}_k(z) = \frac{\tilde{Y}_k(z)}{X(z)} = \sum_{l=0}^{K-1} (z \exp\{-j2\pi k/K\})^{-l} E_l(z^K) \quad (4.24)$$

Since  $(z \exp\{-j2\pi k/K\})^K = z^K$ , evaluating Equation 4.24 at  $z = z \exp\{-j2\pi k/K\}$ , yields

$$H_k(z) = H_0(z \exp\{-j2\pi k/K\}) \quad (4.25)$$

This shows  $\tilde{y}_k[m]$  is the output of a bandpass filter centered at  $\omega = 2\pi k/K$ .

In order to reduce the speed at which the filter, IDFT matrix and modulators operate, the R-fold decimators can be translated from the output of the modulators to the input of the polyphase filters as shown in Figure 4.3. Letting  $\tilde{v}_k^d[m]$  be the decimated output of



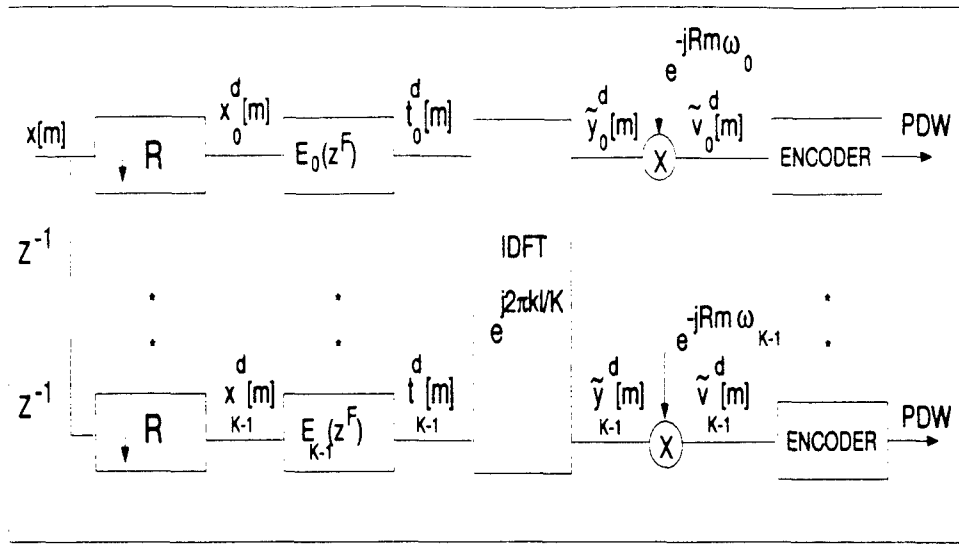


Figure 4.3 Nonmaximally Decimated Polyphase UDFT Filter Bank

the  $k^{th}$  channel, then

$$\tilde{v}_k^d[m] = \tilde{y}_k^d[Rm]e^{-j\omega_k Rm} = \sum_{l=0}^{K-1} t_l[Rm]e^{-j2\pi kl/K}e^{-j\omega_k Rm} \quad (4.26)$$

Since the first exponential term is independent of  $R$ , the decimators can be moved to the inputs of the IDFT matrix to allow calculation of the IDFT and modulators at the decimated rate. Decimating the output of the  $l^{th}$  polyphase filter gives the sequence  $s_l^d[m] = t_l[Rm]$  so that the  $z$ -transform of  $s_l^d[m]$  becomes

$$S_l(z) = \frac{1}{R} \sum_{r=0}^{R-1} E_l \left( z^{K/R} \exp \left\{ \frac{-j2\pi Kr}{R} \right\} \right) X \left( z^{1/R} \exp \left\{ \frac{-j2\pi r}{R} \right\} \right) z^{-l/R} \exp \left\{ \frac{j2\pi rl}{R} \right\} \quad (4.27)$$

Since the ratio  $K/R = F$  is an integer and  $\exp\{\frac{-j2\pi Kr}{R}\} = 1$  for all integers  $r$  then

$$S_l(z) = E_l(z^F) \frac{1}{R} \sum_{r=0}^{R-1} X \left( z^{1/R} \exp \left\{ \frac{-j2\pi r}{R} \right\} \right) \left( z^{-l/R} \exp \left\{ \frac{j2\pi rl}{R} \right\} \right) \quad (4.28)$$

This is equivalent to replacing the filter  $E_l(z^K)$  with  $E_l(z^F)$  and moving the  $R$  fold decimator to the front of the filter bank as shown in Figure 4.3. All operations in the filter bank are now accomplished at  $1/R^{th}$  the rate of the input data.

In terms of the time domain, the output of the  $l^{th}$  polyphase component can be written as

$$e_l[n] = \begin{cases} h_o[Kn/F + l], & \text{for } n = pF \text{ for integer } p \\ 0, & \text{otherwise} \end{cases} \quad (4.29)$$

Since  $N = KN_o$ , the decimated output of each polyphase component becomes

$$t_l^d[m] = \sum_{n=0}^{N_o-1} h_o[l + mK]x[Rm - l - nK] \quad (4.30)$$

The IDFT output is then

$$\tilde{y}_k[m] = \sum_{l=0}^{K-1} \sum_{n=0}^{N_o-1} h_o[l + nK]x[Rm - l - nK]e^{j2\pi kl/K} \quad (4.31)$$

This expression is equivalent to Equation 4.4 for the decimated case where  $m = Rm$ .

The unique implementation of a nonmaximally decimated polyphase filter bank for  $K = FR = 32$  consists of zero-padding each polyphase component filter with  $F - 1$  zeros and decimating by  $R$  the input across all  $K$  channels prior to taking the IDFT. For the real data model, since  $R < .5K$ , the maximum rate of decimation is 8. Alternatively, for the complex data model, since  $R < K$ , the maximum rate of decimation is 16. Thus, if the accuracy of the frequency estimates obtained from the complex data model are equal to those obtained from the real data model, the complex data model should be employed.

As shown in Appendix I, this architecture for performing channelization via a non-maximally decimated polyphase filter bank is mathematically related to a similar method used by Rabiner and Crochiere (55). However, the architecture derived here requires only  $N$  unique filter coefficients to produce the equivalent STFT. Conversely, the architecture proposed by Rabiner and Crochiere (55) requires the installation of  $K$ ,  $F$ -fold expanders following decimators and uses  $F$  subsets of  $R$  unique filters; each of length  $N/R$ , for a total of  $FN$  filter coefficients to produce the equivalent STFT. This architectural design is an original contribution of this research and has been accepted for publication (99). In addition, the architecture has been assigned Serial Number 08/816,951 by the U.S. Patent and Trademark Office and given a filing date of 21 Jan 97.

### 4.3 Parameter Estimation

**4.3.1 Estimation Accuracy.** As described in Chapter III, a critical step in the development and analysis of point estimators is to accurately model the PDF of the data. For both the real and complex cases, the data model was shown to be equivalent to viewing  $\mathbf{y}$  as a set of observations generated from the deterministic system  $\Lambda\mathbf{b}$  corrupted by the zero mean, normally distributed, correlated noise vector  $\mathbf{w}$ .

$$\mathbf{y} = \Lambda_i \mathbf{b} + \mathbf{w} \quad (4.32)$$

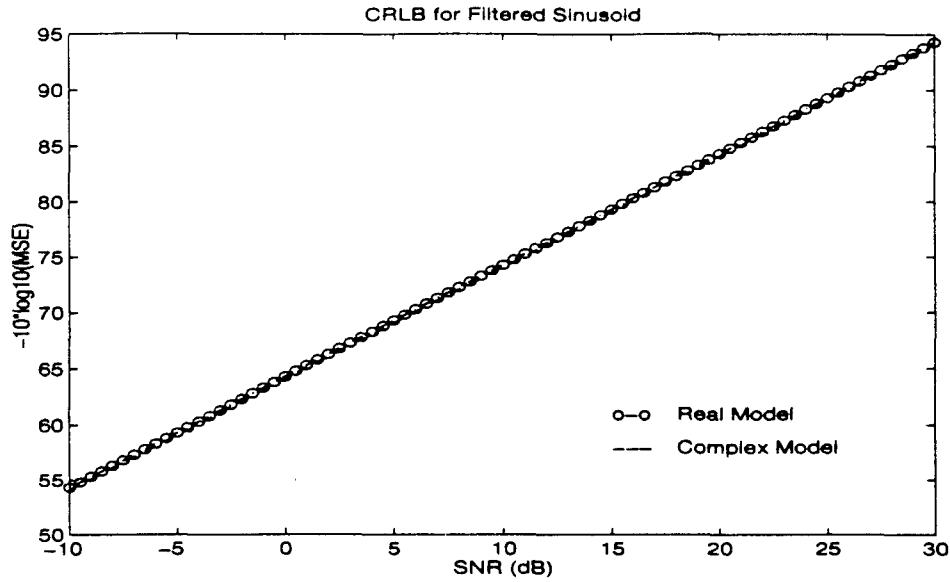
where  $i = \mathcal{R}$  denotes the real data model and  $i = \mathcal{C}$  denotes the complex data model. The problem of estimating the frequencies, phases, and amplitudes was changed to the problem of estimating the signal matrix,  $\Lambda_i$ , and the scaling coefficient vector,  $\mathbf{b}$ , from the measurements in  $\mathbf{y}$ . In addition, for both models, the PDF of the measurement vector had the form

$$p(\mathbf{y}; \boldsymbol{\theta}) = [2\pi]^{-M_i} |\mathbf{K}_w|^{-.5} \exp\left[-\frac{1}{2}(\mathbf{y} - \Lambda_i \mathbf{b})^T \mathbf{K}_w^{-1} (\mathbf{y} - \Lambda_i \mathbf{b})\right] \quad (4.33)$$

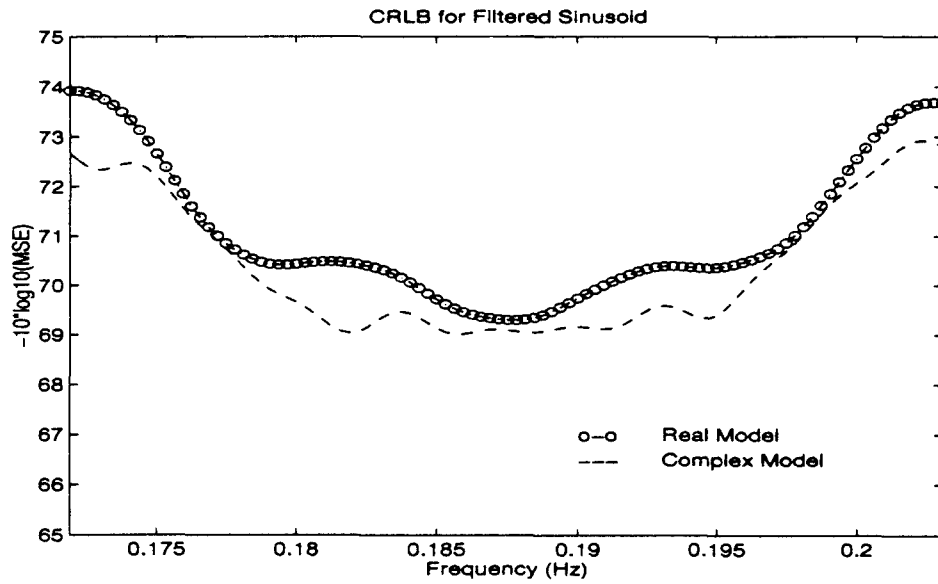
where  $M_i = -.5M$  for the real model and  $M_i = M$  for the complex model. For both models, as shown in Chapter III, the CRLB, which provides a limit on the accuracy of any unbiased parameter estimator, can be calculated from

$$\mathbf{K}_{\boldsymbol{\theta}} - \left[ \frac{\partial \mathbf{m}_y}{\partial \theta_i} \right]^T [\mathbf{K}_w]^{-1} \left[ \frac{\partial \mathbf{m}_y}{\partial \theta_j} \right] \geq [0] \quad (4.34)$$

Figure 4.4 shows the CRLB for the frequency of a single sinusoid using both the real and complex models for data obtained from channel six prior to decimation. As shown in the figure, the CRLB for the frequency of a sinusoid within the filter passband is nearly the same for either model. Thus, there should be virtually no loss in estimation accuracy attributable to the use of the complex data model instead of the real data model. However, since the complex data can be decimated at twice the rate of the real data without aliasing, the complex data model of Equation 4.19 will be employed to develop point estimates of the sinusoidal parameters.



(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .1875, \phi_1^o = \pi/3]$ .



(b) Sinusoid Parameters:  $[b_1^o = 1, \phi_1^o = \pi/3]$  and  $SNR = 5\text{dB}$ .

Figure 4.4 **CRLB for Real and Complex Data Models:** Variance bounded as function of frequency and SNR over filter passband for undecimated signal of block length  $M = 32$ . Filter characteristics: Center Frequency:  $3/16$ ; Bandwidth:  $1/32$ ; Length:  $N = 192$ .

**4.3.2 ML Estimation.** As shown in Chapter III, parameter estimators, whose estimates of the sinusoidal parameters,  $\theta = [b^T, f^T]^T$ , attain the CRLB, can be developed by maximizing the parameter likelihood function. For the complex data model, the ML estimate for  $\theta$  can be obtained by minimizing the following objective function, with respect to the parameters in  $\theta$ :

$$J(\theta) = [y - \Lambda_c b]^T K_w^{-1} [y - \Lambda_c b] \quad (4.35)$$

For a fixed ML estimate of  $f$ , the ML estimator for  $b$  is unique and found as (22), (40)

$$\hat{b}(ML) = [\Lambda_c^T K_w^{-1} \Lambda_c]^{-1} \Lambda_c^T K_w^{-1} y \quad (4.36)$$

The ML estimate for  $f$ , on the other hand, minimizes the following objective function (40):

$$J(f) = y^T K_w^{-1} y - y^T K_w^{-1} \Lambda_c [\Lambda_c^T K_w^{-1} \Lambda_c]^{-1} \Lambda_c^T K_w^{-1} y \quad (4.37)$$

Again, as shown in Chapter III, this objective function, termed the ML frequency objective function, is a highly nonlinear function with respect to  $f$ . Direct minimization involved a computationally intensive search in the  $P$  dimensional frequency domain based on gradient search algorithms. In general, though the method of gradient search can provide accurate frequency estimates, this technique usually requires too many iterations to converge, and is too sensitive to the initial estimates for this EW application. Alternatively, a simpler estimator for the frequencies can be developed by employing LP modeling techniques for the complex data model.

**4.3.3 LP Modeling.** In general, a complex signal,  $\tilde{s}[m]$ , consisting of the sum of  $P$  complex sinusoids can be represented as the sum of  $P$  complex modes (73:484-485)

$$\tilde{s}[m] = \sum_{k=1}^P [b_k e^{j\phi_k}] e^{j\omega_k m} = \sum_{k=1}^P [b_k e^{j\phi_k}] z_k^m = s_r[m] + j s_i[m] \quad (4.38)$$

Here,  $z_k = e^{j2\pi f_k}$  while  $s_r[m]$  and  $s_i[m]$  are the real and imaginary components of  $\tilde{s}[m]$ . A  $P^{th}$  order complex polynomial,  $A_o(z)$ , which incorporates these modes can be constructed

as

$$A_o(z) = \tilde{a}_o[0] \prod_{k=1}^{P_o} (1 - z_k z^{-1}) = \sum_{p=0}^{P_o} \tilde{a}[p] z^{-p} \quad (4.39)$$

where the LP coefficients,  $\tilde{a}[p]$ , are complex quantities given by  $\tilde{a}[p] = a_r[p] + j a_i[p]$ . Thus,  $A(e^{j\omega_k m}) = 0$  for each of the  $P$  frequencies present in  $\tilde{s}[m]$ . Since each mode is a zero of  $A(z)$ , the  $P$  frequencies present in  $\tilde{s}[m]$  can be found as roots to  $A(z)$  and the corresponding time domain expression becomes the complex LP equation.

$$\sum_{p=0}^P \tilde{a}[p] \tilde{s}[m-p] = 0 \quad (4.40)$$

For this equation to be true for  $M$  data samples,  $0 \leq m \leq M-1$ , both the real and imaginary components must be zero. That is

$$\begin{aligned} \sum_{p=0}^P a_r[p] s_r[m-p] - a_i[p] s_i[m-p] &= 0 \\ \sum_{p=0}^P a_r[p] s_i[m-p] + a_i[p] s_r[m-p] &= 0 \end{aligned} \quad (4.41)$$

To construct a matrix form of this equation for  $m = P \dots M-1$ , define the LP coefficient vectors,  $\mathbf{a}_r$  and  $\mathbf{a}_i$ , as  $\mathbf{a}_r = [a_r[0] \dots a_r[P]]^T$  and  $\mathbf{a}_i = [a_i[0] \dots a_i[P]]^T$  and the component signal matrices,  $S_r$  and  $S_i$ , as

$$[S_r]_{k,l} = s_r[M+1-k-l] \quad (4.42)$$

$$[S_i]_{k,l} = s_i[M+1-k-l] \quad (4.43)$$

for  $k = 1 \dots M-P$  and  $l = 1 \dots P$ . Using these definitions, the complex LP general linear model becomes

$$S = \begin{bmatrix} S_r & -S_i \\ S_i & S_r \end{bmatrix} \begin{bmatrix} \mathbf{a}_r \\ \mathbf{a}_i \end{bmatrix} = S \mathbf{a}_c = \mathbf{0} \quad (4.44)$$

From the complex data model, since  $\tilde{y}[m] = \tilde{s}[m] + \tilde{w}[m]$ , in terms of the measurements, the complex LP general linear model can be written as

$$Y a_c a = \begin{bmatrix} Y_r & -Y_i \\ Y_i & Y_r \end{bmatrix} \begin{bmatrix} a_r \\ a_i \end{bmatrix} = \begin{bmatrix} W_r & -W_i \\ W_i & W_r \end{bmatrix} \begin{bmatrix} a_r \\ a_i \end{bmatrix} = \begin{bmatrix} z_r \\ z_i \end{bmatrix} = z \quad (4.45)$$

Here, the matrices,  $Y_r$ ,  $Y_i$ ,  $W_i$  and  $W_r$  are defined in a manner similar to  $S_r$  and  $S_i$  while  $z_r$  and  $z_i$  can be considered as the real and imaginary error vectors. Finally, define  $A_r$  as the  $M - P$  by  $M$  matrix of real coefficients

$$[A_r]_{k,l} = \begin{cases} a_r[k-l] & \text{for } l \leq k \leq l+P \\ 0 & \text{otherwise} \end{cases} \quad (4.46)$$

and  $A_i$  as the associated matrix of imaginary coefficients

$$[A_i]_{k,l} = \begin{cases} a_i[k-l] & \text{for } l \leq k \leq l+P \\ 0 & \text{otherwise} \end{cases} \quad (4.47)$$

for  $l = 1 \dots M - P$ . The vector of errors,  $z$ , can be written as

$$z = \begin{bmatrix} z_r \\ z_i \end{bmatrix} = \begin{bmatrix} A_R^T & -A_I^T \\ A_I^T & A_R^T \end{bmatrix} \begin{bmatrix} w_r \\ w_i \end{bmatrix} = A_c^T w \quad (4.48)$$

Here,  $A_c$  is the matrix of coefficients for the complex LP model. As a linear transformation of  $w$ , a zero mean, normally distributed random vector,  $z$  becomes a zero mean, normally distributed random vector with covariance matrix

$$K_z = A_c^T K_w A_c \quad (4.49)$$

Premultiplying Equation 4.45 by  $G(a_c)$ , the Cholesky decomposition of  $K_z^{-1}$ , yields a new form of the complex LP general linear model

$$G(a_c) Y a_c = Y(a_c) a_c = e(a_c) \quad (4.50)$$

Here,  $\mathbf{e}(\mathbf{a}_c)$  is a zero mean, normally distributed random vector with covariance matrix  $K_{\mathbf{e}(\mathbf{a}_c)} = I$ . The estimation of the  $P$  frequencies in  $\tilde{s}[m]$  has been recast as the estimation of the  $2P + 2$  LP coefficients using the LP general linear model derived in Equation 4.50. The number of LP coefficients to be estimated can be reduced further by imposing sinusoidal constraints on the coefficients.

As shown in Appendix C, for  $P$  complex sinusoids, the LP coefficients are complex symmetric with  $\tilde{a}[p] = \tilde{a}[P-p]^*$ . Thus only  $P+1$  coefficients are necessary to determine  $\mathbf{a}_c$  completely. As in Chapter III, let  $\alpha$  be the set of  $P+1$  real coefficients,  $\alpha = [\alpha[0] \dots \alpha[P]]^T$ . With  $P$  being an even number, the symmetry constraints can be imposed by defining the vector  $\alpha$  as

$$\begin{aligned}\alpha[0] &= a_R[L] \\ \alpha[k+1] &= a_R[k] \quad \text{for } k = 0 \dots L-1 \\ \alpha[L+l+1] &= a_I[l] \quad \text{for } l = 0 \dots L-1\end{aligned}\tag{4.51}$$

where  $L = P/2$ . Alternatively, if  $P$  is an odd number, let  $L = (P-1)/2$ . The symmetry constraints can be imposed by defining the vector  $\alpha$  as

$$\begin{aligned}\alpha[0] &= a_R[L] \\ \alpha[k+1] &= a_R[k] \quad \text{for } k = 0 \dots L-1 \\ \alpha[L+l+1] &= a_I[l] \quad \text{for } l = 0 \dots L\end{aligned}\tag{4.52}$$

In either case, the LP coefficient vector  $\mathbf{a}_c^T = [\mathbf{a}_r^T, \mathbf{a}_i^T]$  can be derived from  $\alpha$  as  $\mathbf{a}_c = B\alpha$  where  $B$  is the matrix relating the  $P+1$  unique coefficients in  $\alpha$  to  $\mathbf{a}_c$ .

Imposing these constraints on the complex LP general linear model yields

$$G(\mathbf{a}_c)YB\alpha = G(\alpha)Y_C\alpha = \mathbf{e}(\alpha)\tag{4.53}$$

The search for the LP coefficients, contained in  $\mathbf{a}_c$ , has been reduced to the identification of the  $P+1$  constrained coefficients contained in  $\alpha$ . Here,  $Y_C$  is an  $M-2P$  by  $P+1$  matrix of rank  $P+1$ . Now let  $\mathbf{y}_c$  be the first column of  $Y_C$  and let  $Y_c$  be an  $M-2P$  by  $P$  matrix of the remaining columns of  $Y_C$ . By constraining  $\alpha[0] = 1$  and defining  $\mathbf{y}_c(\alpha) = G(\alpha)\mathbf{y}_c$



and  $Y_c(\alpha) = G(\alpha)Y_c$ , a final form of the complex LP general linear model becomes

$$\mathbf{y}_c(\alpha) = -Y_c(\alpha)\alpha_o + \mathbf{e}(\alpha) \quad (4.54)$$

Equation 4.54 represents the true general linear model relating the LP coefficients to the frequencies present in  $P$  complex sinusoids in noise and is an original contribution of this research and has been submitted for publication (97). The problem now is to estimate the  $P$  unique LP parameters in  $\alpha_o$  given the  $M$  observations with  $M > P$ .

**4.3.4 LP Model-ML Frequency Estimation Relationship.** As described in Chapter III, there exists a relationship between the complex LP general linear model of Equation 4.54 and the ML frequency objective function of Equation 4.37. First, assume that  $\mathbf{y}$  has length  $2M > 2P$  and  $\Lambda_c$  has  $2P$  linearly independent columns. From Equation 4.19, since  $\Lambda_c$  is an  $2M$  by  $2P$  matrix with  $2P$  linearly independent columns, the column vectors of  $\Lambda_c$ , denoted  $\lambda_i$ , for  $i = 1 \dots 2P$ , span a  $2M$  dimensional subspace,  $V_1$ , of the real vectors,  $R^{2M}$ . As such, there exists a  $2M - 2P$  dimensional orthogonal subspace, denoted  $V_2$ , which is spanned by the vectors,  $\mathbf{a}_j$  for  $j = 1 \dots 2M - 2P$ , such that if  $\psi \in V_1$  and  $\gamma \in V_2$  then  $\psi^T \gamma = 0$  (83:132-152). Now consider the matrix  $A_c$  given by Equation 4.48. From Equation 4.41, for the complex sinusoids,  $A_c^T \Lambda_c = [0]$ . The  $2M - 2P$  columns of  $A$  are orthogonal to the  $2P$  columns of  $\Lambda_c$ . Furthermore, since the columns of  $A_c$  are linearly independent, they form a basis for  $V_2$ , the subspace orthogonal to  $V_1$ . Employing the same steps as in Chapter III, minimization of  $J(\mathbf{f})$  with respect to  $\mathbf{f}$  can be written as a minimization of  $J(A_c)$  with respect to the basis vectors in  $A_c$  where

$$J(A_c) = \mathbf{y}^T A_c (A_c^T K_w A_c)^{-1} A_c^T \mathbf{y} \quad (4.55)$$

Finally, since each column of  $A_c$  is derived from the LP coefficient vector,  $\mathbf{a}_c$ , then  $J(A_c)$  can be written as

$$J(\mathbf{a}_c) = \mathbf{a}_c^T Y^T (A_c^T K_w A_c)^{-1} Y \mathbf{a}_c \quad (4.56)$$

Employing the complex sinusoidal constraints via the matrix  $B$  and taking the Cholesky decomposition the inverse of  $A_C^T K_w A_C$  yields

$$J(\alpha) = \alpha^T Y_C^T G(\alpha)^T G(\alpha) Y_C \alpha \quad (4.57)$$

By defining the real and imaginary error vectors as  $e_r(\alpha) = [e_r[M-1] \dots e_r[P]]^T$  and  $e_i(\alpha) = [e_i[M-1] \dots e_i[P]]^T$  and constraining  $\alpha[0] = 1$  yields

$$e(\alpha) = G(\alpha) Y_C \alpha = y_c(\alpha) + Y_c(\alpha) \alpha_o \quad (4.58)$$

where  $e(\alpha)^T = [e_r(\alpha)^T, e_i(\alpha)^T]$  represents the concatenation of the real and imaginary error vectors. This equation is identical to the complex LP general linear model derived in Chapter III. By properly incorporating the effects of the filter coefficients and the LP coefficients in developing the LP general linear model, this derivation shows the exact relationship between the LP coefficients and ML complex frequency estimates. The set of LP coefficients which minimize the squared error defined by the LP general linear model will provide ML estimates of the frequencies. Furthermore, via the invariance of ML estimators, the LP coefficients which minimize the squared error are the ML estimates of the coefficients.

**4.3.5 LP Coefficient Estimation.** Ideally, the optimal estimator for the LP coefficients would minimize  $J(\alpha)$  so that the resulting frequencies would be ML estimates. As discussed in Chapter III, minimization of  $J(\alpha)$  with respect to  $\alpha$  is a nonlinear optimization problem. Direct minimization involved a computationally intensive search in the  $P$  dimensional coefficient domain based on gradient descent algorithms. In general, though the method of gradient descent can provide accurate LP coefficient, and consequently, frequency estimates, this technique usually requires too many iterations to converge to an answer and is too sensitive to the initial estimates for this EW application.

Alternatively, since Equation 4.54, has the same form as the LP general linear model derived in Chapter III, a fixed point estimator can be constructed to estimate  $\alpha$  from the complex data. As shown in Chapter III, the ILS and ITLS fixed point estimators were

found to produce estimates nearly as accurate as those attained by the direct minimization of  $J(\alpha)$ . However, these estimators took fewer iterations to converge and were less sensitive to initial estimates. Since the ILS estimator was found to produce estimates as least as accurate as the ITLS estimator while being less computationally intense, this estimator will be used to provide frequency estimates of the decimated output of the EW receiver.

From the system model, with  $\alpha_o^0 = \mathbf{0}_P$ , let  $\alpha_o^i$  be the  $i^{\text{th}}$  estimate of the LP coefficients. Employing the ILS fixed point algorithm of Chapter III yields

$$\alpha_o^{i+1} = - [Y_c^T(\alpha_o^i) Y_c(\alpha_o^i)]^{-1} Y_c^T(\alpha_o^i) \mathbf{y}(\alpha_o^i) \quad (4.59)$$

When  $\alpha_o^{i+1} = \alpha_o^i$ , the error vector,  $e(\alpha^i)$  is orthogonal to the columns of  $Y_c(\alpha^i)$  and  $\alpha_o^i = \mathcal{L}_{ILS}(\alpha_o^i)$ . A fixed point has been reached and the estimate of the LP coefficients becomes  $\hat{\alpha}_o = \alpha_o^i$ .

Figure 4.5 and Figure 4.6 show the accuracy of the true and the approximated ILS estimator, in relation to the frequency estimation technique employed in the EW receiver, for estimating the frequency of a single sinusoid at the filter output. As shown in these figures, the ILS estimator attains the CRLB at an SNR approximately 5dB lower than the current phase-based WPA estimator and maintains this accuracy across the passband of the filter. In addition, the approximate ILS estimator, constructed with  $K_w = \sigma^2 I$ , performs almost as well as the true ILS estimator. This indicates the effect of the filter and decimation serve to decorrelate the data enough that the approximate model should be adequate for estimating the frequency of a single filter sinusoid.

Figure 4.7 and Figure 4.8 show the accuracy of the true and the approximated ILS estimator for estimating the frequency of a two sinusoids at the filter output. At the present time, the EW receiver does not have the capability to estimate the frequencies of two time-coincident sinusoids within the filter passband. As shown in Figure 4.7a, for two sinusoids with frequencies near the passband frequencies, the frequency estimates do not reach the CRLB at low SNRs. This is probably due to the severe reduction in signal amplitude occurring at the edge of the filter. Figure 4.7b shows the estimation accuracy for two closely spaced sinusoids. Here, the decimated frequency separation of  $df = .01$  can

not be resolved through the use of a Fourier Transform. The approximate ILS estimator appears to outperform the true ILS estimator at low SNRs. Close analysis of the figure shows the approximate ILS estimator is providing biased frequency estimates at these low SNRs. Above the threshold of about 1dB, both estimators perform about the same. Finally, Figure 4.8c shows the accuracy of both ILS estimators for a more typical set of signals. Here, since the two sinusoids are within the passband and are adequately separated in frequency, both estimators attained the CRLB at less than -5dB. As with the single sinusoid, the approximate ILS estimator performed as well as the true ILS estimator. Thus, the approximate estimator should be adequate for estimating the frequency of two filtered sinusoids.

**4.3.6 Complex Model Approximation.** For this particular prototype filter,  $h_o[n]$ , and decimation rate,  $R = 16$ , experimental analysis shows  $H_i^T H_r \approx [0]$  and  $H_r^T H_r \approx H_i^T H_i$ . Using these approximations, as shown in Appendix H, the PDF of the complex data model can be written in the form of a complex multivariate Gaussian PDF (40:505-507)

$$p(\tilde{\mathbf{y}}, \boldsymbol{\theta}) = [\pi]^{-M} |\tilde{\mathbf{K}}_{\tilde{\mathbf{w}}}|^{-1} \exp \left\{ -(\tilde{\mathbf{y}} - \tilde{\mathbf{\Lambda}} \tilde{\mathbf{b}})^H \tilde{\mathbf{K}}_{\tilde{\mathbf{w}}}^{-1} (\tilde{\mathbf{y}} - \tilde{\mathbf{\Lambda}} \tilde{\mathbf{b}}) \right\} \quad (4.60)$$

Here, the  $M$  by  $P$  complex matrix,  $\tilde{\mathbf{\Lambda}}$  is defined by  $[\tilde{\mathbf{\Lambda}}]_{lp} = e^{j2\pi R f_p [M-l+N-1]}$  and the complex vector  $\tilde{\mathbf{b}}$  is defined by  $\tilde{b}_p = b_p e^{j\phi_p}$ . In addition,  $\tilde{\mathbf{K}}_{\tilde{\mathbf{w}}} = \sigma^2 \tilde{\mathbf{H}}^H \tilde{\mathbf{H}}$  where  $\tilde{\mathbf{H}}$  is a complex  $R(M-1)+N$  by  $M$  filter matrix constructed using the complex filter coefficients in Equation 4.8. This is equivalent to viewing  $\tilde{\mathbf{y}}[m]$  as the output of a complex filter  $\{\tilde{h}[n]\}$  due to a linear sum of  $P$  complex sinusoids in normally distributed, independent, complex noise,  $\tilde{\eta}[m]$ . Using this approximation, equations for the input and output data become

$$\tilde{\mathbf{x}}[m] = \sum_{p=1}^P b_p^o e^{j\phi_p^o} e^{j2\pi f_p m} + \tilde{\eta}[m] \quad \text{and} \quad \tilde{\mathbf{y}}[m] = \sum_{p=1}^P b_p e^{j\phi_p} e^{j2\pi f_p m} + \tilde{\mathbf{w}}[m] \quad (4.61)$$

where  $\tilde{\mathbf{w}}[m]$  is a zero-mean, complex Gaussian noise sequence with a covariance matrix given by  $\mathbf{K}_{\tilde{\mathbf{w}}} = \sigma^2 \tilde{\mathbf{H}}^H \tilde{\mathbf{H}}$ . This approximation allows a simpler method of characterizing the output of the digital receiver and is an original contribution of this research.

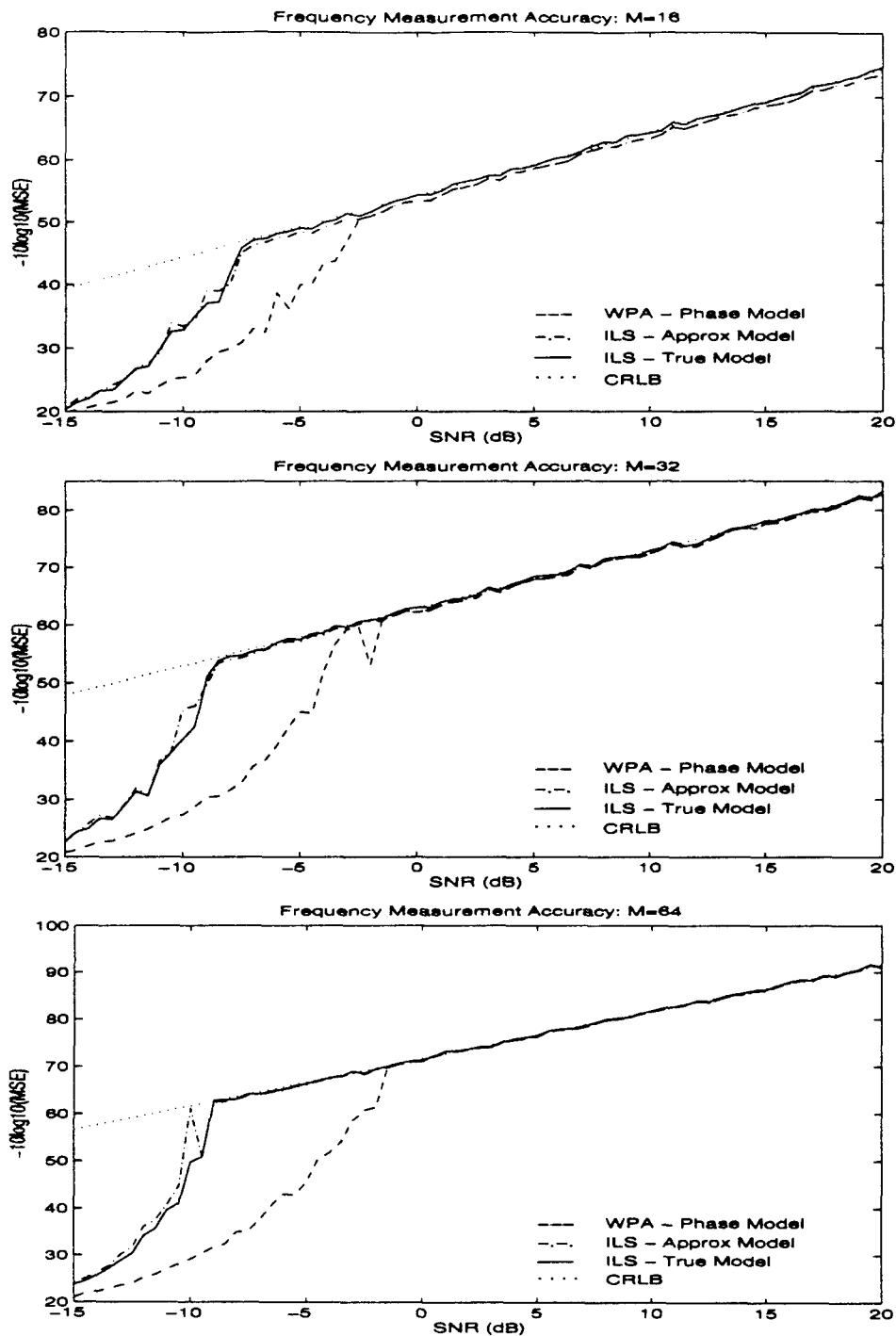
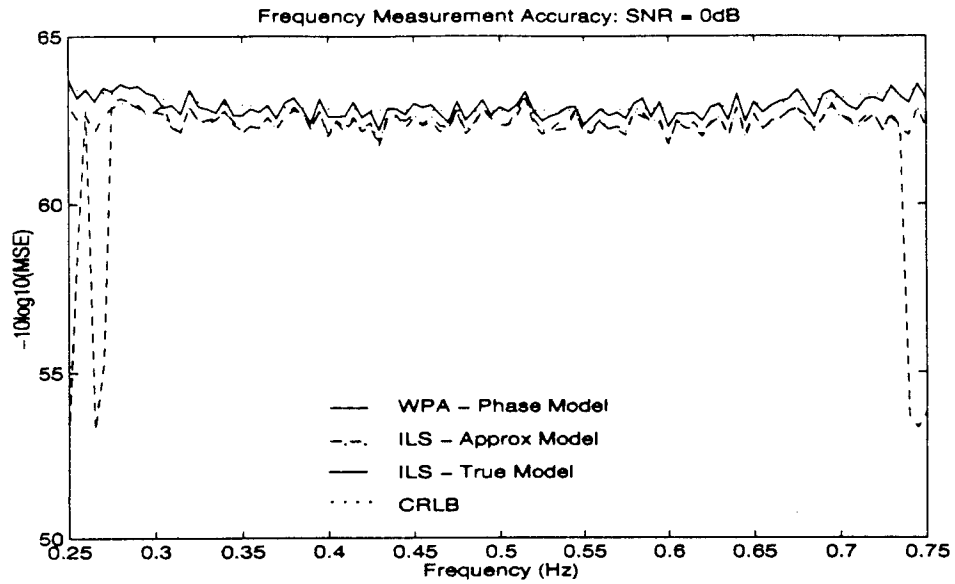
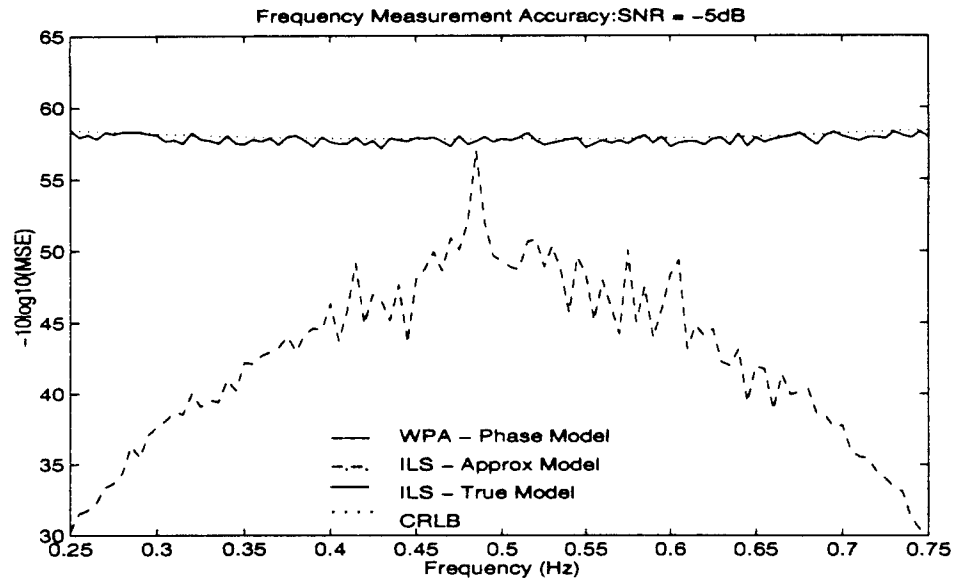


Figure 4.5 **ILS Estimation Accuracy for Single Sinusoid:** MSE calculated from 500 independent realizations and  $M = 16, 32, 64$  samples. Ten iterations allowed for convergence. The incorrect model assumed  $K_w = \sigma^2 I$ . Correct model used  $K_w$ . Sinusoid Parameters:  $[b_1^o = 1, f_1 = .0235, \phi_1^o = 3\pi/8]$ ;  $\text{SNR} = -10\log_{10}(2\sigma^2)$ ;  $\text{MSE} = \frac{1}{500} \sum_{i=1}^{500} (\hat{f}_1[i] - f_1)$ .

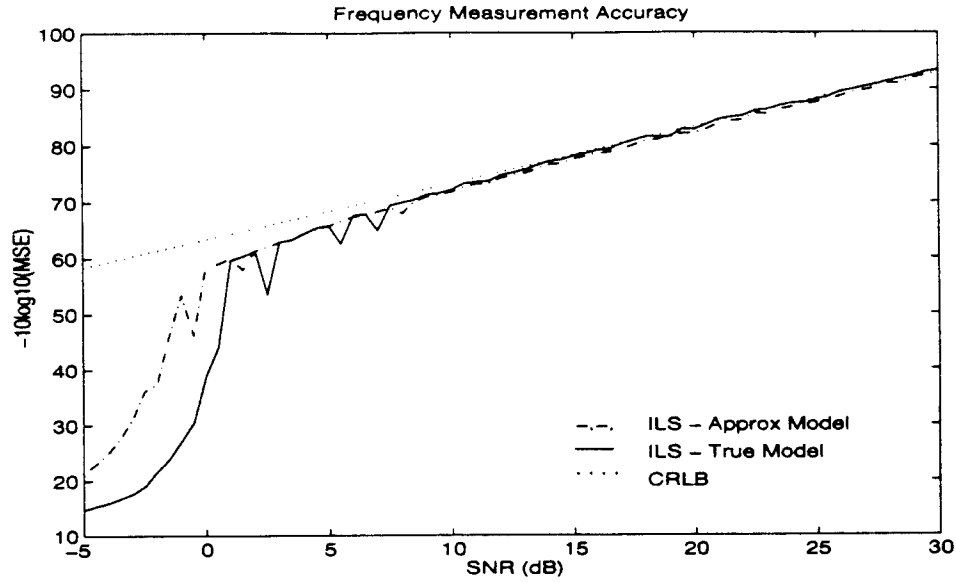


(a) SNR = 0 dB.

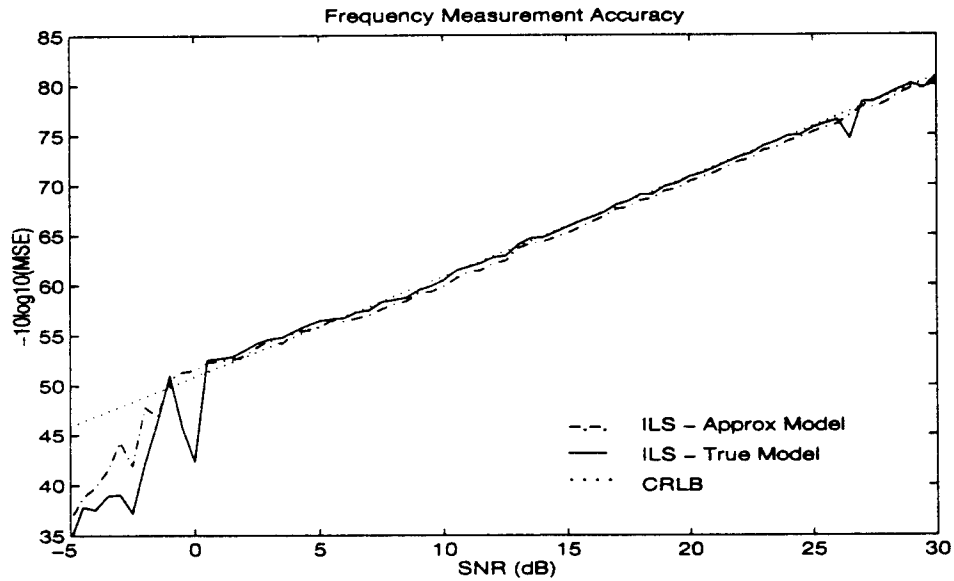


(b) SNR = -5 dB.

Figure 4.6 **ILS Frequency Estimation Accuracy Across Bandwidth:** MSE calculated from 500 independent realizations of block length  $M = 32$ . Ten iterations allowed for convergence. The approximate model assumed  $K_w = \sigma^2 I$  while the true model used correct  $K_w$ .  $SNR = -10 \log_{10}(2\sigma^2)$ ;  $MSE = \frac{1}{500} \sum_{i=1}^{500} (\hat{f}_1[i] - f_1)$  Sinusoidal Parameters:  $[b_1^0 = 1, \phi_1^0 = 3\pi/8]$ .



(a) Parameters:  $[b_1^o = 1, f_1 = .0164, \phi_1^o = 3\pi/8]$ ;  $[b_2^o = 1, f_2 = .0461, \phi_2^o = 4\pi/3]$ .



(b) Parameters:  $[b_1^o = 1, f_1 = .0309, \phi_1^o = 3\pi/8]$ ;  $[b_2^o = 1, f_2 = .0316, \phi_2^o = 4\pi/3]$ .

Figure 4.7 **ILS Frequency Estimation Accuracy for Two Difficult Sinusoids:** MSE calculated from 500 independent realizations of block length  $M = 32$ . Ten iterations allowed for convergence. The approximate model assumed  $K_w = \sigma^2 I$  while the true model used  $K_w$ .  $SNR = -10\log_{10}(2\sigma^2)$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p[i] - f_p)$ .

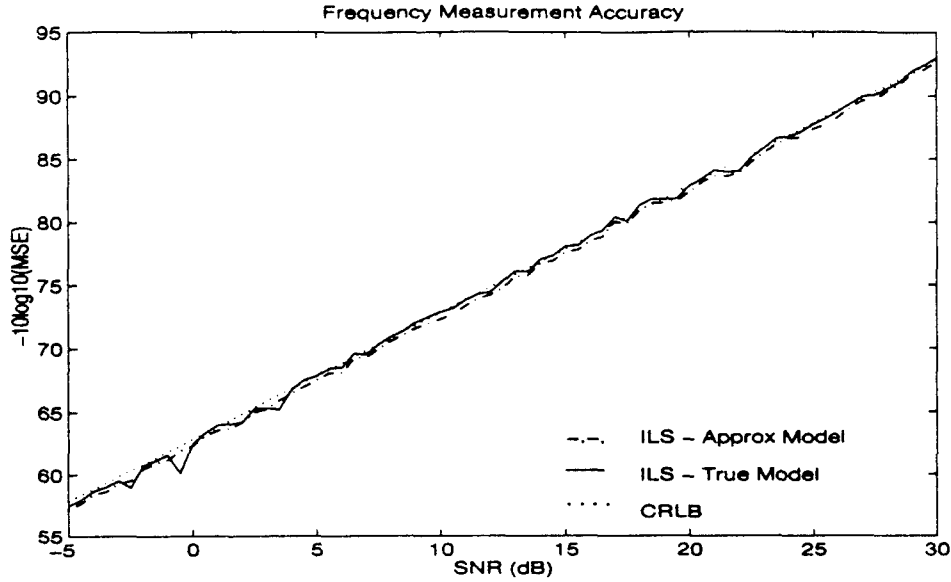


Figure 4.8 ILS Frequency Estimation Accuracy for Two Typical Sinusoids: MSE calculated from 500 independent realizations of block length  $M = 32$ . Ten iterations allowed for convergence. The approximate model assumed  $K_w = \sigma^2 I$  while the true model used  $K_w$ .  $SNR = -10 \log_{10}(2\sigma^2)$ ;  $MSE = \frac{1}{1000} \sum_{p=1}^2 \sum_{i=1}^{500} (\hat{f}_p[i] - f_p)$ .

4.3.7 *Bounding the Estimation Error.* As shown in Chapter III, the error in a point estimate of the LP coefficient and, consequently, the frequencies can be bounded if the variance in the estimates is known or can be calculated. From Equation 3.165, since the complex LP general linear model has the same form as the real LP general linear model, an unbiased point estimate of the variance can be calculated directly from the measurements as

$$\hat{K}_{\hat{\alpha}_o} = [Y_c^T G^T(\hat{\alpha}_o) G(\hat{\alpha}_o) Y_c]^{-1} \quad (4.62)$$

Figure 4.9 shows the average accuracy of the point estimates of the LP coefficient variance for one and two filtered sinusoids. These results show the point estimates provided by Equation 4.62 are, on average, very close to the true values of the variance.

The accuracy of the variance estimates indicates that confidence intervals for the frequencies can be established directly from the variance estimates using Equation 3.148 from Chapter III. That is, once  $\hat{K}_{\hat{\alpha}_o}$  has been estimated, the variance in the frequency



estimate can be calculated from

$$\hat{V}\{\hat{f}_p\} = \left[ \frac{g_p(\alpha_o)}{\partial \alpha_o} \right]^T \hat{K} \alpha_o \left[ \frac{g_p(\alpha_o)}{\partial \alpha_o} \right] \quad (4.63)$$

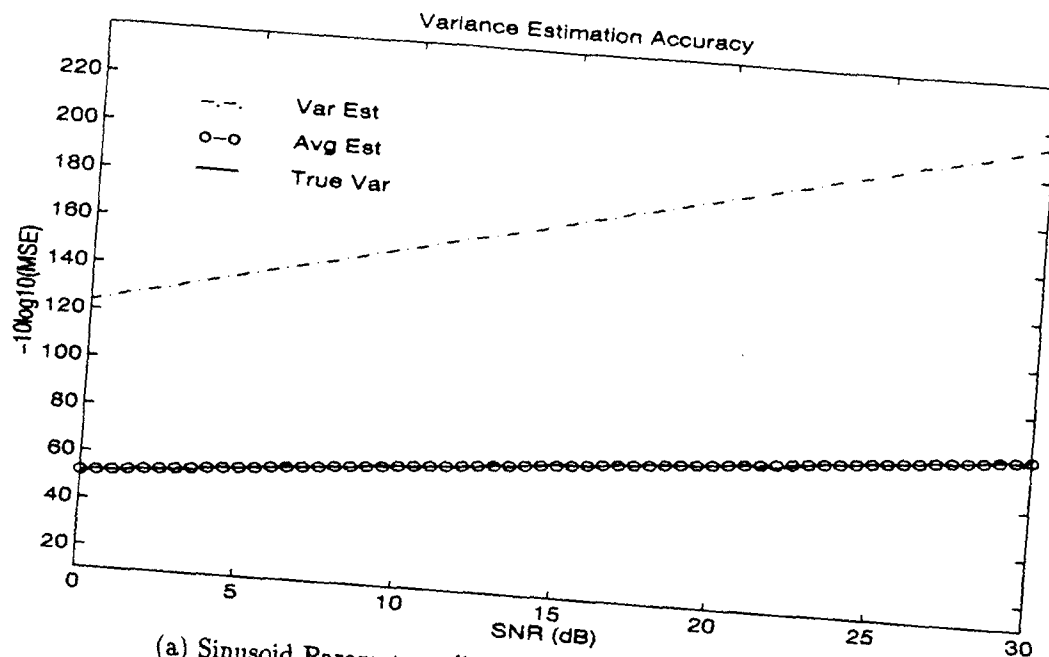
where  $g_p()$  is the function relating the LP coefficients directly to the frequencies. In particular, for a single complex sinusoid derived from the constrained LP coefficients,  $g_p()$  and  $V\{\}$  take the form

$$\hat{f}_1 = -\frac{1}{\pi} \tan^{-1}\{\hat{a}_i[0]\}; \quad V\{f_1\} \approx \pi^{-2}(1 + \hat{a}_i[0]^2)^{-2} \hat{V}\{a_i[0]\} \quad (4.64)$$

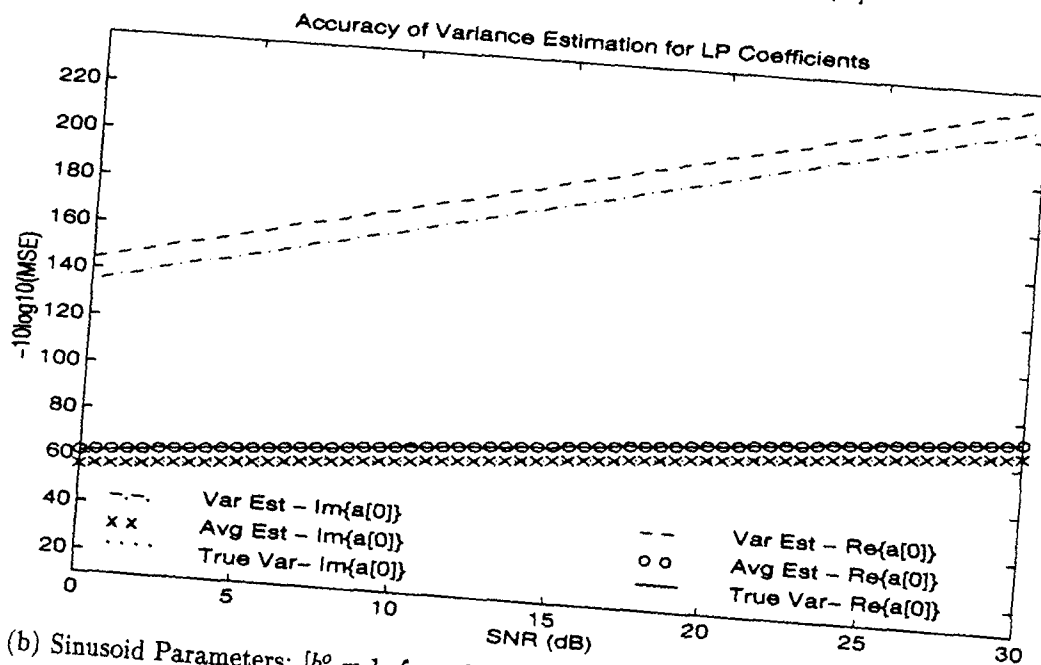
Figure 4.10 shows results of using this equation to establish confidence intervals for the frequency of a single sinusoid in the passband of filter. Thus, from the point estimates of the frequency,  $\hat{f}_1$ , and standard deviation,  $\hat{\sigma}_f$ , the true frequency was within one, two and three standard deviations of the estimate about 65%, 95% and 99% of the time.

For an EW receiver, knowledge of the estimation accuracy can improve the accuracy of the classifications. Currently, the signal descriptions employed by typical EW classification systems consist of range limit descriptions for the signal parameters. These range limits are usually constructed with the 'worst-case' measurement assumption of a receiver operating at threshold. This assumption forces the classification system to extend the range limits of the signal models which induces regions in which the ranges overlap within the models. This overlap in ranges, in turn, causes multiple identifications to be made for a single signal. Providing an assessment of the goodness of an estimate, along with the point estimate itself, can lead to improved classification accuracies by allowing the width of the classification range limits to be compressed or expanded to match the actual measurement accuracy. Finally, providing an estimate of the 'goodness' of the measurement will allow poor measurements to be discarded prior to classification.

**4.3.8 Simplified Single Sinusoidal Estimator.** Examination of the simulations shows the effect of decimation by  $R = 16$  is to decorrelate the noise such that the LS estimates of  $\tilde{\alpha}$  for  $\tilde{K}_{\tilde{w}} \approx \sigma_2 I$  are as accurate as those attained using the true noise



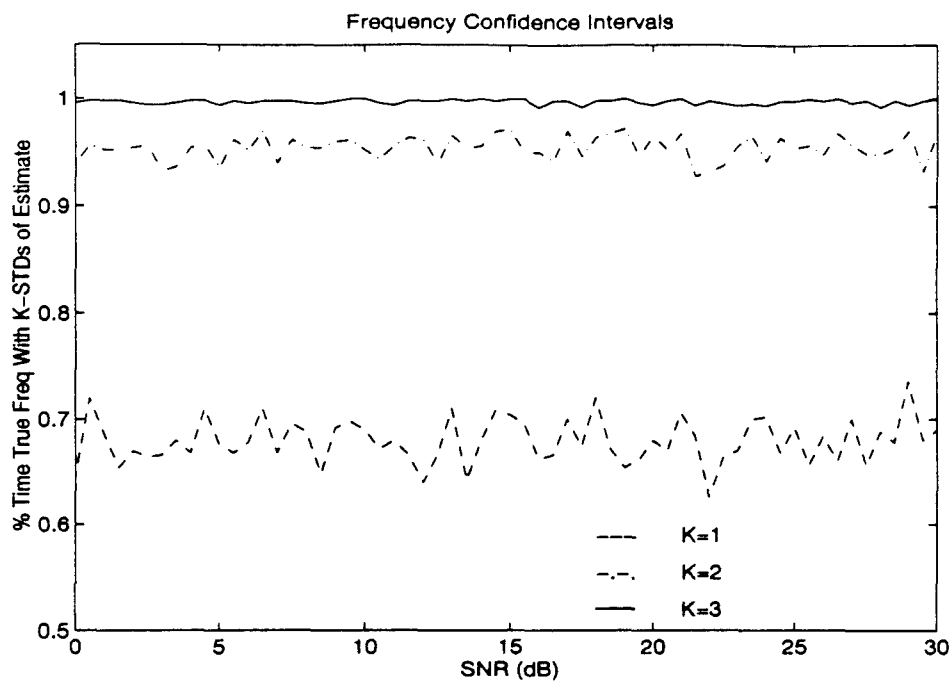
(a) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .0234, \phi_1^o = \pi/3]$ .



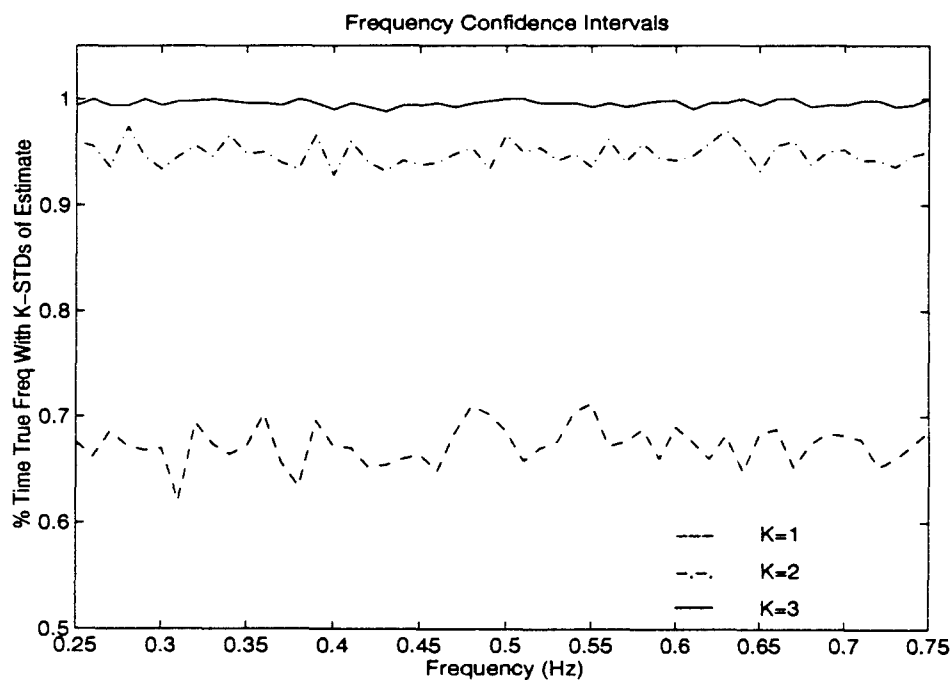
(b) Sinusoid Parameters:  $[b_1^o = 1, f_1 = .0297, \phi_1^o = \pi/3], [b_2^o = 1, f_2 = .0375, \phi_2^o = 4\pi/3]$ .

**Figure 4.9 Variance Estimation of Complex LP Coefficients:** Variance estimated at SNR intervals of .5dB from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. LP coefficients estimated by ILS algorithm. Ten iterations allowed convergence. True variance found from CRLB. SNR calculated as  $SNR = -10 \log_{10}(2\sigma^2)$ . Avg Est =  $\frac{1}{500} \sum_{i=1}^{500} \hat{V}\{a[p]\}_i$ .  

$$\text{Var Est} = \frac{1}{500} \sum_{i=1}^{500} [\hat{V}\{a[p]\}_i - V\{a[p]\}_{TRUE}]^2$$



a: Parameters:  $[b_1 = 1, f_1 = .0234, \phi_1 = \pi/3]$  and SNR = 5dB.



b: Parameters:  $[b_1 = 1, \phi_1 = \pi/3]$  and SNR = -5dB.

Figure 4.10 **Frequency Confidence Intervals:** Confidence intervals established at SNR intervals of .5dB and frequency intervals .005 from 500 independent realizations of  $M = 32$  samples of sinusoids in noise. LP coefficients estimated by ILS algorithm. Ten iterations allowed convergence. Confidence interval:  $|\hat{f}_p(i) - f_p| < k\hat{\sigma}_f$  for  $k = 1, 2, 3$ . SNR calculated as  $SNR = -10\log_{10}(2\sigma^2)$ .

covariance. This approximation will allow an extremely simple form of the LS estimator to be implemented for a single sinusoid.

From Equation 4.40, with  $P = 1$  and  $\tilde{a}[0] = 1$  the LS fixed point estimator for  $\tilde{a}[1] = \tilde{a} = -e^{j2\pi f_1}$ , can be written as

$$\tilde{a}_{i+1} = -\frac{\tilde{\mathbf{y}}_1^H (\tilde{\mathbf{A}}_i^H \tilde{\mathbf{A}}_i)^{-1} \tilde{\mathbf{y}}_0}{\tilde{\mathbf{y}}_1^H (\tilde{\mathbf{A}}_i^H \tilde{\mathbf{A}}_i)^{-1} \tilde{\mathbf{y}}_1} \quad (4.65)$$

Here  $\tilde{\mathbf{y}}_1 = [\tilde{y}[M-1] \dots \tilde{y}[1]]^T$  and  $\tilde{\mathbf{y}}_0 = [\tilde{y}[M-2] \dots \tilde{y}[0]]^T$  while  $\tilde{\mathbf{A}}$  is the  $M-P$  by  $M$  matrix of LP coefficients defined in a manner similar to the  $\mathbf{A}$  matrix of Chapter III. Examination of the matrix product shows  $\tilde{\mathbf{A}}_i^H \tilde{\mathbf{A}}_i$  has the following characteristics:

$$[\tilde{\mathbf{A}}_i^H \tilde{\mathbf{A}}_i]_{kl} = \begin{cases} 2 & \text{if } k = l \\ \tilde{a}_i & \text{if } l = k + 1 \\ \tilde{a}_i^* & \text{if } k - 1 = l \\ 0 & \text{otherwise} \end{cases} \quad (4.66)$$

Using a Cholesky decomposition,  $\tilde{\mathbf{A}}_i^H \tilde{\mathbf{A}}_i = \tilde{\mathbf{R}}_i \tilde{\mathbf{R}}_i^H$  where  $\tilde{\mathbf{R}}_i$  is an  $M-1$  by  $M-1$  lower triangular matrix given by

$$\tilde{r}_{mn} = [\tilde{\mathbf{R}}_i]_{mn} = \begin{cases} \sqrt{\frac{m+1}{m}} & \text{if } m = n \\ \tilde{a}_i^* / \tilde{r}_{(m-1), (m-1)} & \text{if } n = m - 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.67)$$

Expansion yields  $[\tilde{\mathbf{R}}_i^{-1}]_{mn} = [\tilde{a}_i^*]^{m-n} c_{mn}$  where the real quantity,  $c_{mn}$ , is zero for  $m > n$  and

$$c_{mn} = \frac{(-1)^{(m-n)}}{\left[ \left[ \prod_{i=n}^{m-1} \tilde{r}_{ii}^2 \right] \tilde{r}_{mm} \right]} \quad (4.68)$$

for  $n \leq m$ . Thus,  $c_{mn}$  is a known constant derived only from the real constants defined in  $\tilde{r}_{mm}$ . Defining the  $M-1$  by  $M-1$  matrix  $\mathbf{D}$  of constants so that  $d_{kl} = [\mathbf{D}]_{kl} = \sum_{n=1}^{M-1} c_{nk} c_{nl}$ ,

then the matrix inverse calculations can be written as

$$[\tilde{A}_i^H \tilde{A}_i]_{kl}^{-1} = \sum_{n=1}^{M-1} [\tilde{R}_i^H]_{kn}^{-1} [\tilde{R}_i]_{nl}^{-1} = [\tilde{a}_i^*]^k d_{kl} [\tilde{a}_i]^l \quad (4.69)$$

Thus, by constructing the vectors  $\tilde{v}_i$  and  $\tilde{u}_i$  from  $\tilde{y}_1$  and  $\tilde{y}_2$  as  $\{\tilde{v}_i\}_m = \tilde{a}_i^m \{\tilde{y}_1\}_m$  and  $\{\tilde{u}_i\}_m = \tilde{a}_i^m \{\tilde{y}_2\}_m$ , a simplified estimator for  $\tilde{a}_i$  can now be written as

$$\tilde{a}_i = -\frac{\tilde{y}_1^H (\tilde{A}_i^H \tilde{A}_i)^{-1} \tilde{y}_2}{\tilde{y}_1^H (\tilde{A}_i^H \tilde{A}_i)^{-1} \tilde{y}_1} = -\frac{\tilde{v}_i^H D \tilde{u}_i}{\tilde{v}_i^H D \tilde{v}_i} \quad (4.70)$$

The matrix inverse operation has been replaced by a simple vector-matrix-vector product operation. This efficient representation is useful when the number of data samples,  $M$ , to be processed becomes large. The development of this simplified estimator for a single complex sinusoid is an original contribution of this research and has been submitted for publication (96).

#### 4.4 Summary

In this chapter, the ILS fixed point estimation technique was successfully used to estimate the frequencies of sinusoids at the output of the digital EW receiver being developed by AFRL/WL (16). After discussing the general receiver architecture, models were derived for the real and complex data output from bandpass filters and represent an original contribution of this dissertation. The model derived for the complex data shows the effective decimation rate of the output data be doubled over that of the real model without a loss of frequency estimation accuracy.

A new method to efficiently implement the receiver as a nonmaximally decimated, UDFT polyphase filter bank is then derived based on the complex data model. This derivation has been accepted for publication and has been assigned Serial Number 08/816,951 by the U.S. Patent and Trademark Office (99).

Based on this architecture and the complex data model, a complex form of the LP general linear model is developed for estimating the frequencies of  $P$  real sinusoids

residing in the filter passband. This new form of the LP general linear model is an original contribution of this research and has been submitted for publication.

Identification of the LP coefficients is then shown to be related to ML frequency estimation and an ILS fixed point estimator is derived for estimating the coefficients. For one sinusoid, simulations indicate the ILS estimator attains the CRLB for frequency estimation at SNRs significantly lower than those currently employed in the receiver. The ILS estimator was then successfully used to estimate the frequencies of two time-coincident sinusoids within the filter passband; a capability which currently does not exist in the receiver.

This chapter concluded by showing the PDF of the complex data can be approximately modeled as a complex Gaussian PDF. Using this approximation, a simplified frequency estimator of a single sinusoid was derived which effectively removed the requirement to invert an  $M - 1$  square complex matrix. This new single frequency estimator for complex sinusoids is an original contribution of this dissertation and is currently under review for publication.

Finally, in addition to providing the Air Force with a complete probabilistic characterization of their next generation digital EW receiver system, this dissertation derived a new method of estimating the frequencies of filtered sinusoids. The frequency estimators developed as a result of this research will greatly extend the operational envelope of the EW receiver.

## V. Conclusion

### 5.1 Introduction

The dissertation investigated the problem of estimating the parameters of real and complex filtered sinusoids in noise. In Chapter III, the general theory of parameter estimation for real filtered sinusoids in noise was developed and used to construct accurate parameter estimators. The resulting analysis, models, and estimators derived in this chapter have applications across the spectrum of applied science, engineering and statistics. In Chapter IV, the results of this research were applied to the Air Force's next-generation EW receiver. The resulting analysis, models, and estimators derived in this chapter led to an efficient hardware implementation of the receiver while significantly improving the receiver's operational capability.

### 5.2 Theoretical Contributions

As shown in Chapter III, the first contribution of this research was a mathematical model, given by Equation 3.13, for any system required to estimate the amplitudes, phases and frequencies of real, filtered sinusoids. Specifically, this dissertation showed that the steady state output of a filter, due to an input consisting of a linear sum of  $P$  real sinusoids in zero-mean, independent, normally distributed noise, can be modeled as the output from a deterministic system,  $\Lambda\mathbf{b}$ , corrupted by zero-mean, normally distributed noise,  $\mathbf{w}$ , with a covariance matrix,  $K_{\mathbf{w}}$ , determined by the filter characteristics. Estimation of the sinusoidal parameters was recast as the estimation of a deterministic system from the set of measurements,  $\mathbf{y}$ , embedded in colored noise and a bound on the parameter estimation accuracy was derived.

Using this model, a new set of ML estimators for the scaling parameters,  $\mathbf{b}$ , and frequencies,  $\mathbf{f}$ , were derived which correctly account for the correlation in the noise due to the effects of the filter. This dissertation showed the ML estimator for  $\mathbf{b}$ , given in Equation 3.23, can be found via a simple linear regression on the measurement vector  $\mathbf{y}$  once an ML estimate of the frequencies,  $\mathbf{f}$ , is made. The ML estimate of the frequencies, on the other hand, is independent of  $\mathbf{b}$  and can be found by optimizing the ML frequency objec-

tive function,  $J(\mathbf{f})$ , of Equation 3.25. Since this objective function is a highly nonlinear function with respect to  $\mathbf{f}$ , the ML frequency estimator derived in Equation 3.30 involved a computationally intensive search in the  $P$  dimensional frequency based on gradient search algorithms. Simulations indicate the accuracy of these ML estimators will achieve the theoretical bounds provided the correlation in the noise, due to the filter, is correctly modeled. Failure to incorporate the filter effects in the ML model leads to suboptimal parameter estimates. The main drawback in using these estimators for an EW receiver is that the gradient search algorithm employed to provide ML frequency estimates usually required too many iterations to converge and relied on an accurate initial estimate.

To overcome these drawbacks, the problem of frequency estimation was recast as the problem of estimating the coefficients of the  $2P^{th}$  order LP polynomial. After describing the relationship between the LP coefficients and the sinusoids, a linear model, parameterized by the LP coefficients, was derived in Equation 3.53. This model, which incorporates the effects of both the filter and the coefficients upon the noise, is the true linear model relating the LP coefficients to the filtered sinusoids in noise and is an original contribution of this research (98).

Based on this general linear model, two estimators, based on fixed point theory, were then derived for estimating the LP coefficients and represent another original contribution of this research. The ILS estimator, as given by Equation 3.84, was based on an iterative least squares solution to an over-determined system of equations. The ITLS estimator, as given by Equation 3.100, was based on an iterative total least squares solution. Simulations indicate both methods provide near-minimum variance, unbiased estimates of the LP coefficients, and consequently, the sinusoidal frequencies, over a wide range of SNRs. Furthermore, each algorithm usually took less than ten iterations to converge and was relatively insensitive to the initial estimate. In addition, the initial coefficient estimate can be obtained directly within the framework of the algorithms; no additional estimation routines are required to provide the initial estimates. All other factors being equal, the ILS method should be preferred over the ITLS method since the ITLS method requires an SVD of an  $M - 2P$  by  $P + 1$  matrix at each iteration.



This dissertation then derived the exact relationship between ML frequency estimation of the filtered sinusoids and ML estimation of the LP coefficients. Specifically, the set of LP coefficients which minimize the square error objective function,  $J(\alpha)$ , defined in Equation 3.124, was shown to provide ML frequency estimates and vice versa. This original contribution led directly to the development of the IEGD estimator, given by Equation 3.127, for minimizing  $J(\alpha)$ . This estimator provides ML estimates of the LP coefficients and, consequently, the frequencies of filtered sinusoids, by employing an iterative exact gradient descent algorithm to minimize  $J(\alpha)$ . Simulations indicated the estimates found via the IEGD algorithm do minimize the LP objective function thus providing ML estimates of the LP coefficients and, consequently, the frequencies. However, this algorithm usually required too many iterations to converge and was too sensitive to the initial estimates to be of much use as a real-time frequency estimator.

This dissertation then proved the equivalence between the IGLS algorithm, previously presented as a method of minimizing the LP objective function,  $J(\alpha)$ , and the ILS fixed point estimator of Equation 3.84. This proof is an original contribution of this research for it correctly casts the IGLS algorithm as a fixed point estimator, not a minimization algorithm.

In addition, the IQML algorithm was proven to be exactly equivalent to the ITLS fixed point estimator of Equation 3.100. This is another original contribution of this dissertation for it correctly casts the IQML algorithm, widely accepted as the premier method of minimizing  $J(\alpha)$ , as a fixed point estimator, not a minimization algorithm. In addition, for the simulations completed, the ILS/IGLS algorithm actually produced, on average, more accurate frequency estimates than the more complicated ITLS/IQML algorithm. The IQML algorithm requires an eigenvalue decomposition of a  $P + 1$  by  $P + 1$  square matrix at each iteration whereas the IGLS algorithm only requires the inversion of a  $P$  square matrix. Thus, if the small degradation in performance over the IEGD can be tolerated, the ILS/IGLS estimator should be preferred over the ITLS/IQML estimator for providing point estimates of the LP coefficients and, consequently, the sinusoidal frequencies.

Finally, a method was developed for bounding the estimation error of the LP coefficient estimates, and consequently, the frequency estimates, based strictly on one realization

of the measurement vector. Specifically, as shown by Equation 3.165, this dissertation proved the point estimate of the covariance matrix of the LP coefficients is unbiased. This derivation is an original contribution of this research and allows the estimate of the LP coefficient variance to be transformed into an unbiased estimate of the frequency covariance matrix. In addition, as shown in Equation 3.147 and Equation 3.148, from a single estimate of the variance of the LP coefficients and the frequencies, the estimation error can be bounded and confidence intervals constructed. Simulations indicate these estimates of the variance can be used to gauge the accuracy of the point estimates of both the LP coefficients and the frequencies. This gauge can then be employed to develop confidence intervals to aid any decision making process based on a single set of measurements.

### 5.3 *Applied Contributions*

In this dissertation, a new architecture, assigned Serial Number 08/816,951 by the U.S. Patent and Trademark Office, was derived to implement a nonmaximally decimated UDFT polyphase filter bank. This architecture is based on a new data model relating parameter estimation of real sinusoids to the parameter estimation of complex sinusoids and is an original contribution of this research (99). Specifically, as shown in Figure 4.3, this new data model transforms the estimation of the parameters of real sinusoids in real colored noise to the estimation of the parameters of complex sinusoids in complex colored noise. Estimation of the real sinusoidal parameters is recast as the estimation of the deterministic linear system,  $\Lambda_c \mathbf{b}$ , from a set of complex measurements,  $\mathbf{y}$  embedded in colored noise. Simulations for the EW architecture indicate this transformation can be attained with negligible loss in estimation accuracy attributable to the complex model.

In addition, this new data model shows how complex data, originating from a sum of real sinusoids in real noise passing through a complex bandpass filter, can be decimated at twice the rate of the real data without a loss in parameter estimation accuracy. This improvement in the decimation rate, exploited by transferring the decimators to input of the filterbank as shown in Figure 4.3, allows a significant reduction in the speed at which the filters, IDFT matrix, modulator and parameter encoder must operate.

Based on this new model for a complex representation of real sinusoids, a new set of ML estimators for the scaling parameters,  $\mathbf{b}$ , and frequencies,  $\mathbf{f}$ , were derived which correctly account for the correlation in the noise due to the effects of the filter. This dissertation showed the ML estimator for  $\mathbf{b}$ , given in Equation 4.36, can be found via a simple linear regression on the measurement vector  $\mathbf{y}$  once an ML estimate of  $\mathbf{f}$  is made. As with real sinusoids, the ML estimate of the frequencies, on the other hand, is independent of  $\mathbf{b}$  and can be found by optimizing the ML frequency objective function,  $J(\mathbf{f})$  of Equation 4.37 .

Since this objective function, is a highly nonlinear function with respect to  $\mathbf{f}$ , the problem of frequency estimation was recast as the problem of estimating the coefficients of the  $P^{th}$  order complex LP polynomial given by Equation 4.40. After describing the relationship between the LP coefficients and the complex sinusoids, a linear model, parameterized by the LP coefficients, was derived in Equation 4.54. This model, which incorporates the effects of both the filter and the coefficients upon the noise, is the true linear model relating the LP coefficients to complex filtered sinusoids in noise and is an original contribution of this research (97).

Identification of the LP coefficients was then shown to be related to ML frequency estimation and an ILS fixed point estimator was derived for estimating the coefficients. For one sinusoid, simulations indicate the ILS estimator attains the CRLB for frequency estimation at SNRs significantly lower than those attained by the estimation algorithm currently employed in the receiver. The ILS estimator was then successfully used to estimate the frequencies of two time-coincident sinusoids within the filter passband; a capability which currently does not exist in the receiver.

Finally, the results of this applied research showed the PDF of the complex data out of a given channel of the EW receiver filterbank, can be approximately modeled as a complex Gaussian PDF. Using this approximation, a simplified frequency estimator for a single complex sinusoid was derived which effectively removed the requirement to invert an  $M - 1$  square complex matrix. This new single frequency estimator for complex sinusoids is an original contribution of this dissertation and has been submitted for review (96).

#### 5.4 *Summary*

The results of this research, whether taken individually or collectively, represent a major contribution to the theory of signal processing and parameter estimation. In particular, this research builds the bridge connecting sinusoidal frequency estimation with LP linear system modeling and derives the connection between real and complex sinusoidal parameter estimation. In addition, the analysis, models and estimators constructed as a result of this connection will significantly improve the operational envelope of the Air Force's next generation EW receiver.

## Appendix A. Vector-Matrix Differentiation

This appendix derives the identities involving the differentiation of matrix products.

### A.1 Definitions

Let  $\theta$  be  $P$  dimensional vector with  $\theta = [\theta_1 \dots \theta_P]^T$ . Let  $\mathbf{x}$  be  $Q$  dimensional vector with  $\mathbf{x} = [x_1 \dots x_Q]^T$  where each element of  $\mathbf{x}$  is a function of  $\theta$ . That is  $x_q = g_q(\theta)$ . The first and second derivatives of  $\mathbf{x}$  with respect to  $\theta_i$  and  $\theta_j$  are defined as

$$\left[ \frac{\partial \mathbf{x}}{\partial \theta_i} \right]_q = \frac{\partial x_q}{\partial \theta_i}; \quad \left[ \frac{\partial^2 \mathbf{x}}{\partial \theta_i \partial \theta_j} \right]_q = \frac{\partial^2 x_q}{\partial \theta_i \partial \theta_j} \quad (\text{A.1})$$

Let  $U$  be a  $K$  by  $L$  dimensional matrix where each element of  $U$  is a function of  $\theta$ . That is  $u_{k,l} = g_{k,l}(\theta)$ . The first and second partial derivatives of  $U$  with respect to  $\theta_i$  and  $\theta_j$  are defined as

$$\left[ \frac{\partial U}{\partial \theta_i} \right]_{k,l} = \frac{\partial u_{k,l}}{\partial \theta_i}; \quad \left[ \frac{\partial^2 U}{\partial \theta_i \partial \theta_j} \right]_{k,l} = \frac{\partial^2 u_{k,l}}{\partial \theta_i \partial \theta_j} \quad (\text{A.2})$$

### A.2 Matrix-Vector Product Derivative

Let  $U$  be an  $M$  by  $N$  matrix and  $\mathbf{x}$  by an  $N$  dimensional vector. Let  $u_{m,n}$  and  $x_n$  be functions of the variables  $\theta_i$  and  $\theta_j$ . In addition, let  $\mathbf{z} = U\mathbf{x}$ . Then

$$\frac{\partial z_m}{\partial \theta_i} = \sum_{n=1}^N \frac{\partial u_{m,n}}{\partial \theta_i} x_n + u_{m,n} \frac{\partial x_n}{\partial \theta_i}; \quad \frac{\partial \mathbf{z}}{\partial \theta_i} = \left[ \frac{\partial U}{\partial \theta_i} \right] \mathbf{x} + U \left[ \frac{\partial \mathbf{x}}{\partial \theta_i} \right] \quad (\text{A.3})$$

while the second partial derivatives become

$$\frac{\partial^2 z_m}{\partial \theta_i \partial \theta_j} = \sum_{n=1}^N \frac{\partial^2 u_{m,n}}{\partial \theta_i \partial \theta_j} x_n + u_{m,n} \frac{\partial^2 x_n}{\partial \theta_i \partial \theta_j} + \frac{\partial u_{m,n}}{\partial \theta_i} \frac{\partial x_n}{\partial \theta_j} + \frac{\partial u_{m,n}}{\partial \theta_j} \frac{\partial x_n}{\partial \theta_i} \quad (\text{A.4})$$

$$\frac{\partial^2 \mathbf{z}}{\partial \theta_i \partial \theta_j} = \left[ \frac{\partial^2 U}{\partial \theta_i \partial \theta_j} \right] \mathbf{x} + U \left[ \frac{\partial^2 \mathbf{x}}{\partial \theta_i \partial \theta_j} \right] + \left[ \frac{\partial U}{\partial \theta_i} \right] \left[ \frac{\partial \mathbf{x}}{\partial \theta_j} \right] + \left[ \frac{\partial U}{\partial \theta_j} \right] \left[ \frac{\partial \mathbf{x}}{\partial \theta_i} \right] \quad (\text{A.5})$$

### A.3 Matrix Product Derivative

Let  $U$  and  $V$  be  $L$  by  $K$  and  $L$  by  $L$  matrices respectively. Let  $u_{k,l}$  and  $v_{l,m}$  be functions of  $\theta$  and define the matrix  $C$  as  $C = U^T V U$ . Then

$$\frac{\partial c_{k,n}}{\partial \theta_p} = \sum_{l=1}^L \sum_{m=1}^L \frac{\partial u_{k,l}}{\partial \theta_p} v_{l,m} u_{m,n} + u_{k,l} \frac{\partial v_{l,m}}{\partial \theta_p} u_{m,n} + u_{k,l} v_{l,m} \frac{\partial u_{m,n}}{\partial \theta_p} \quad (A.6)$$

$$\begin{aligned} \frac{\partial^2 c_{k,n}}{\partial \theta_p \partial \theta_q} = & \sum_{l=1}^L \sum_{m=1}^L \frac{\partial^2 u_{k,l}}{\partial \theta_p \partial \theta_q} v_{l,m} u_{m,n} + \frac{\partial u_{k,l}}{\partial \theta_p} \frac{\partial v_{l,m}}{\partial \theta_q} u_{m,n} + \frac{\partial u_{k,l}}{\partial \theta_p} v_{l,m} \frac{\partial u_{m,n}}{\partial \theta_q} \\ & + \frac{\partial u_{k,l}}{\partial \theta_q} \frac{\partial v_{l,m}}{\partial \theta_p} u_{m,n} + u_{k,l} \frac{\partial^2 v_{l,m}}{\partial \theta_p \partial \theta_q} u_{m,n} + u_{k,l} \frac{\partial v_{l,m}}{\partial \theta_p} \frac{\partial u_{m,n}}{\partial \theta_q} \\ & + \frac{\partial u_{k,l}}{\partial \theta_q} v_{l,m} \frac{\partial u_{m,n}}{\partial \theta_p} + u_{k,l} \frac{\partial v_{l,m}}{\partial \theta_q} \frac{\partial u_{m,n}}{\partial \theta_p} + u_{k,l} v_{l,m} \frac{\partial^2 u_{m,n}}{\partial \theta_p \partial \theta_q} \end{aligned} \quad (A.7)$$

In matrix nomenclature, the partial derivative with respect to  $\theta_p$  is

$$\frac{\partial C}{\partial \theta_p} = \left[ \frac{\partial U}{\partial \theta_p} \right]^T V U + U^T \left[ \frac{\partial V}{\partial \theta_p} \right] U + U^T V \left[ \frac{\partial U}{\partial \theta_p} \right] \quad (A.8)$$

while the partial derivative with respect to  $\theta_p$  and  $\theta_q$  becomes

$$\begin{aligned} \frac{\partial^2 C}{\partial \theta_p \partial \theta_q} = & \left[ \frac{\partial^2 U}{\partial \theta_p \partial \theta_q} \right]^T V U + \left[ \frac{\partial U}{\partial \theta_p} \right]^T \left[ \frac{\partial V}{\partial \theta_q} \right] U + \left[ \frac{\partial U}{\partial \theta_p} \right]^T V \left[ \frac{\partial U}{\partial \theta_q} \right] \\ & + \left[ \frac{\partial U}{\partial \theta_q} \right]^T \left[ \frac{\partial V}{\partial \theta_p} \right] U + U^T \left[ \frac{\partial^2 V}{\partial \theta_p \partial \theta_q} \right] U + U^T \left[ \frac{\partial V}{\partial \theta_p} \right] \left[ \frac{\partial U}{\partial \theta_q} \right] \\ & + \left[ \frac{\partial U}{\partial \theta_q} \right]^T V \left[ \frac{\partial U}{\partial \theta_p} \right] + U^T \left[ \frac{\partial V}{\partial \theta_q} \right] \left[ \frac{\partial U}{\partial \theta_p} \right] + U^T V \left[ \frac{\partial^2 U}{\partial \theta_p \partial \theta_q} \right] \end{aligned} \quad (A.9)$$

### A.4 Inverse Derivative

Let  $C$  be a symmetric, invertible matrix where each element is a function of  $\theta_i$ . Then (40:73-75)

$$\frac{\partial C^{-1}}{\partial \theta_i} = -C^{-1} \left[ \frac{\partial C}{\partial \theta_i} \right] C^{-1} \quad (A.10)$$

and

$$\frac{\partial^2 C^{-1}}{\partial \theta_i \partial \theta_j} = - \left[ \frac{\partial C^{-1}}{\partial \theta_j} \right] \left[ \frac{\partial C}{\partial \theta_i} \right] C^{-1} - C^{-1} \left[ \frac{\partial^2 C}{\partial \theta_i \partial \theta_j} \right] C^{-1} - C^{-1} \left[ \frac{\partial C}{\partial \theta_i} \right] \left[ \frac{\partial C^{-1}}{\partial \theta_j} \right] \quad (A.11)$$

### A.5 Vector-Matrix-Vector Product Derivative

Let  $\mathbf{x}$  be  $K$  dimensional vector and let  $C$  be a  $K$  by  $K$  positive definite matrix. Furthermore, let the scalar  $J$  be defined as  $J = \mathbf{x}^T C^{-1} \mathbf{x}$ . Then

$$\frac{\partial J}{\partial \theta_p} = \sum_{n=1}^K \sum_{m=1}^K \frac{\partial x_m}{\partial \theta_p} C_{m,n}^{-1} x_n + x_m \frac{\partial C_{m,n}^{-1}}{\partial \theta_p} x_n + x_m C_{m,n}^{-1} \frac{\partial x_n}{\partial \theta_p} \quad (\text{A.12})$$

$$\begin{aligned} \frac{\partial^2 J}{\partial \theta_p \partial \theta_q} &= \sum_{n=1}^K \sum_{m=1}^K \frac{\partial^2 x_m}{\partial \theta_p \partial \theta_q} C_{m,n}^{-1} x_n + \frac{\partial x_m}{\partial \theta_p} \frac{\partial C_{m,n}^{-1}}{\partial \theta_q} x_n + \frac{\partial x_m}{\partial \theta_p} C_{m,n}^{-1} \frac{\partial x_n}{\partial \theta_q} \\ &+ \frac{\partial x_m}{\partial \theta_q} \frac{\partial C_{m,n}^{-1}}{\partial \theta_p} x_n + x_m \frac{\partial^2 C_{m,n}^{-1}}{\partial \theta_p \partial \theta_q} x_n + x_m \frac{\partial C_{m,n}^{-1}}{\partial \theta_p} \frac{\partial x_n}{\partial \theta_q} \\ &+ \frac{\partial x_m}{\partial \theta_q} C_{m,n}^{-1} \frac{\partial x_n}{\partial \theta_p} + x_m \frac{\partial C_{m,n}^{-1}}{\partial \theta_q} \frac{\partial x_n}{\partial \theta_p} + x_m C_{m,n}^{-1} \frac{\partial^2 x_n}{\partial \theta_p \partial \theta_q} \end{aligned} \quad (\text{A.13})$$

In vector-matrix notation, the partial derivative with respect to  $\theta_p$  becomes

$$\frac{\partial J}{\partial \theta_p} = 2 \left[ \frac{\partial \mathbf{x}}{\partial \theta_p} \right]^T C^{-1} \mathbf{x} + \mathbf{x}^T \frac{\partial C^{-1}}{\partial \theta_p} \mathbf{x} \quad (\text{A.14})$$

The partial derivative with respect to  $\theta_p$  and  $\theta_q$  can be written as

$$\begin{aligned} \frac{\partial^2 J}{\partial \theta_p \partial \theta_q} &= 2 \left[ \frac{\partial^2 \mathbf{x}}{\partial \theta_p \partial \theta_q} \right]^T C^{-1} \mathbf{x} + 2 \left[ \frac{\partial \mathbf{x}}{\partial \theta_p} \right]^T \left[ \frac{\partial C^{-1}}{\partial \theta_q} \right] \mathbf{x} + 2 \left[ \frac{\partial \mathbf{x}}{\partial \theta_p} \right]^T C^{-1} \left[ \frac{\partial \mathbf{x}}{\partial \theta_q} \right] \\ &+ 2 \left[ \frac{\partial \mathbf{x}}{\partial \theta_q} \right]^T \left[ \frac{\partial C^{-1}}{\partial \theta_p} \right] \mathbf{x} + \mathbf{x}^T \left[ \frac{\partial^2 C^{-1}}{\partial \theta_p \partial \theta_q} \right] \mathbf{x} \end{aligned} \quad (\text{A.15})$$

## Appendix B. Maximizing ML Objective Functions

This appendix derives the scalar and vector identities pertinent to optimization of real objective functions with respect to real and complex vectors. These identities are then used to obtain ML estimators for real and complex data.

### B.1 Real Data ML Maximization

Let  $\mathbf{b} = [b_1 \dots b_Q]^T$  be a vector of  $Q$  real parameters. Let  $\mathbf{x} = [x_1 \dots x_Q]^T$  be a vector of constants and let  $A$  be a square  $Q$  by  $Q$  matrix defined by  $[A]_{kl} = a_{kl}$ . Finally, let  $J$  be a scalar function of  $\mathbf{b}$  so that the gradient of  $J$  with respect to  $\mathbf{b}$  is defined as

$$\nabla_{\mathbf{b}} J = \left[ \frac{\partial J}{\partial b_1} \dots \frac{\partial J}{\partial b_Q} \right]^T \quad (\text{B.1})$$

The results of applying this gradient to vector products and vector-matrix-vector products is given in Table B.2. With  $\boldsymbol{\theta}^T = [\mathbf{b}^T; \mathbf{f}^T]$  and  $\Lambda$  a function of  $\mathbf{f}$  only, the likelihood function is defined as

$$p(\mathbf{y}; \boldsymbol{\theta}) = [2\pi]^{-\frac{M}{2}} |K_{\mathbf{w}}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{y} - \Lambda\mathbf{b})^T K_{\mathbf{w}}^{-1}(\mathbf{y} - \Lambda\mathbf{b})\right] \quad (\text{B.2})$$

Taking the natural logarithm yields the log-likelihood function  $L(\mathbf{y}; \boldsymbol{\theta})$  as

$$L(\mathbf{y}; \boldsymbol{\theta}) = -\frac{M}{2} \ln(2\pi) - \frac{1}{2} \ln |K_{\mathbf{w}}| - \frac{1}{2} [\mathbf{y} - \Lambda\mathbf{b}]^T K_{\mathbf{w}}^{-1} [\mathbf{y} - \Lambda\mathbf{b}] \quad (\text{B.3})$$

Assuming  $K_{\mathbf{w}}$  is known, then maximizing  $L(\mathbf{y}; \boldsymbol{\theta})$  is the same as minimizing

$$J_R(\boldsymbol{\theta}) = [\mathbf{y} - \Lambda\mathbf{b}]^T K_{\mathbf{w}}^{-1} [\mathbf{y} - \Lambda\mathbf{b}] \quad (\text{B.4})$$

Expanding yields

$$J_R(\boldsymbol{\theta}) = \mathbf{y}^T K_{\mathbf{w}}^{-1} \mathbf{y} - 2\mathbf{b}^T \Lambda^T K_{\mathbf{w}}^{-1} \mathbf{y} + \mathbf{b}^T \Lambda^T K_{\mathbf{w}}^{-1} \Lambda \mathbf{b} \quad (\text{B.5})$$



Taking the gradient with respect to  $\mathbf{b}$  yields

$$[\nabla_{\mathbf{b}} J_R(\boldsymbol{\theta})] = -2\Lambda^T K_{\mathbf{w}}^{-1} \mathbf{y} + 2\Lambda^T K_{\mathbf{w}}^{-1} \Lambda \mathbf{b} \quad (\text{B.6})$$

Setting equal to zero and solving for  $\mathbf{b}$  yields the ML estimator which is found via a linear regression on  $\tilde{\mathbf{y}}$  provided  $\mathbf{f}$  is given.

$$\hat{\mathbf{b}}_{ML} = [\Lambda^T K_{\mathbf{w}}^{-1} \Lambda]^{-1} \Lambda^T K_{\mathbf{w}}^{-1} \mathbf{y} \quad (\text{B.7})$$

Substitution of  $\hat{\mathbf{b}}_{ML}$  into  $J_R(\boldsymbol{\theta})$  yields the ML frequency objective function.

$$J_R(\mathbf{f}) = \mathbf{y}^T K_{\mathbf{w}}^{-1} \mathbf{y} - \mathbf{y}^T K_{\mathbf{w}}^{-1} \Lambda [\Lambda^T K_{\mathbf{w}}^{-1} \Lambda]^{-1} \Lambda^T K_{\mathbf{w}}^{-1} \mathbf{y} \quad (\text{B.8})$$

This function must be minimized with respect to  $\mathbf{f}$  to obtain ML estimates. Once  $\hat{\mathbf{f}}_{ML}$  is found, then  $\hat{\mathbf{b}}_{ML}$  is found via Equation B.7.

## B.2 Complex Data ML Maximization

Let the complex number  $\tilde{b}_i$  be defined by real scalars  $b_{iR}$  and  $b_{iI}$  so that  $\tilde{b}_i = b_{iR} + j b_{iI}$  and let  $\tilde{x}_i$  be a complex number defined in a similar fashion. Let  $J$  be a function of a complex variable,  $\tilde{b}_i$ . The goal is to find  $\tilde{b}_i$  so that  $\partial J / \partial \tilde{b}_i = 0$  indicating a stationary point with respect to  $\tilde{b}_i$  (40:517-519) (24:890-894). In general,  $J$  is not an analytic function and cannot be differentiated with respect to  $\tilde{b}_i$  using the normal rules of differential calculus. Instead, the following definition will be applied (40:517)

$$\frac{\partial J}{\partial \tilde{b}_i} = \frac{1}{2} \left[ \frac{\partial J}{\partial b_{iR}} - j \frac{\partial J}{\partial b_{iI}} \right] \quad (\text{B.9})$$

Note, the derivative is zero if and only if the partial derivative of each of the real and imaginary parts is identically zero. Using this definition, scalar functions defined by  $J$  can be minimized over  $b_{iR}$  and  $b_{iI}$  simultaneously (40:515-520). Table B.2 defines the appropriate scalar derivatives used for maximizing the complex ML objective function.

Now, let  $\tilde{\mathbf{b}} = [\tilde{b}_1 \dots \tilde{b}_Q]^T$  be a vector of complex parameters and let  $\tilde{\mathbf{x}} = [\tilde{x}_1 \dots \tilde{x}_Q]^T$  be a vector of complex constants. Furthermore, let  $J(\tilde{\mathbf{b}})$  be any scalar function of  $\tilde{\mathbf{b}}$  and define the gradient of  $J$  with respect to  $\tilde{\mathbf{b}}$  as the vector

$$\nabla_{\tilde{\mathbf{b}}} J = \left[ \frac{\partial J}{\partial \tilde{b}_1} \dots \frac{\partial J}{\partial \tilde{b}_Q} \right]^T \quad (\text{B.10})$$

The results of this gradient, applied to a vector product and a vector matrix product, are given in Table B.2 and these definitions will be used to minimize the ML objective function. With  $\boldsymbol{\theta}^T = [\tilde{\mathbf{b}}^T; \mathbf{f}^T]$  and  $\tilde{\Lambda}$  a function of  $\mathbf{f}$  only, the likelihood function was defined as

$$p(\tilde{\mathbf{y}}; \boldsymbol{\theta}) = [2\pi]^{-M} |\tilde{K}_{\mathbf{w}}|^{-1} \exp[-(\tilde{\mathbf{y}} - \tilde{\Lambda}\tilde{\mathbf{b}})^H \tilde{K}_{\mathbf{w}}^{-1} (\tilde{\mathbf{y}} - \tilde{\Lambda}\tilde{\mathbf{b}})] \quad (\text{B.11})$$

Taking the natural logarithm yields the log-likelihood function  $L(\tilde{\mathbf{y}}; \boldsymbol{\theta})$  as

$$L(\tilde{\mathbf{y}}; \boldsymbol{\theta}) = -M \ln(2\pi) - \ln |\tilde{K}_{\mathbf{w}}| - [\tilde{\mathbf{y}} - \tilde{\Lambda}\tilde{\mathbf{b}}]^H \tilde{K}_{\mathbf{w}}^{-1} [\tilde{\mathbf{y}} - \tilde{\Lambda}\tilde{\mathbf{b}}] \quad (\text{B.12})$$

Assuming  $\tilde{K}_{\mathbf{w}}$  is known, then maximizing  $L(\tilde{\mathbf{y}}; \boldsymbol{\theta})$  is the same as minimizing

$$J_C(\boldsymbol{\theta}) = [\tilde{\mathbf{y}} - \tilde{\Lambda}\tilde{\mathbf{b}}]^H \tilde{K}_{\mathbf{w}}^{-1} [\tilde{\mathbf{y}} - \tilde{\Lambda}\tilde{\mathbf{b}}] \quad (\text{B.13})$$

Expanding yields

$$J_C(\boldsymbol{\theta}) = \tilde{\mathbf{y}}^H \tilde{K}_{\mathbf{w}}^{-1} \tilde{\mathbf{y}} - \tilde{\mathbf{b}}^H \tilde{\Lambda}^H \tilde{K}_{\mathbf{w}}^{-1} \tilde{\mathbf{y}} - \tilde{\mathbf{y}}^H \tilde{K}_{\mathbf{w}}^{-1} \tilde{\Lambda} \tilde{\mathbf{b}} + \tilde{\mathbf{b}}^H \tilde{\Lambda}^H \tilde{K}_{\mathbf{w}}^{-1} \tilde{\Lambda} \tilde{\mathbf{b}} \quad (\text{B.14})$$

Taking the gradient with respect to  $\tilde{\mathbf{b}}$  yields

$$[\nabla_{\tilde{\mathbf{b}}} J(\boldsymbol{\theta})] = -[\tilde{\Lambda}^H \tilde{K}_{\mathbf{w}}^{-1} \tilde{\mathbf{y}}]^* + [\tilde{\Lambda}^H \tilde{K}_{\mathbf{w}}^{-1} \tilde{\Lambda} \tilde{\mathbf{b}}]^* \quad (\text{B.15})$$

Setting equal to zero and solving for  $\tilde{\mathbf{b}}$  yields the ML estimator which is found via a linear regression on  $\tilde{\mathbf{y}}$  provided  $\mathbf{f}$  is given.

Table B.1 Derivatives and Gradients

Real Gradients		
Function	Component Derivative	Vector Gradient
$J = \mathbf{b}^T \mathbf{x} = \sum_{k=1}^Q x_k b_k$	$\frac{\partial J}{\partial b_i} = x_i$	$[\nabla_{\mathbf{b}} J] = \mathbf{x}$
$J = \mathbf{b}^T \mathbf{A} \mathbf{b} = \sum_{k=1}^Q \sum_{l=1}^Q b_k a_{kl} b_l$	$\frac{\partial J}{\partial b_i} = \sum_{k=1}^Q a_{ki} b_k + \sum_{l=1}^Q a_{il} b_l$	$[\nabla_{\mathbf{b}} J] = \mathbf{A} \mathbf{b} + \mathbf{A}^T \mathbf{b}$
Complex Scalar Derivatives		
Function	Component Derivative	Vector Gradient
$J = \tilde{b}_i = b_{iR} + j b_{iI}$	$\frac{\partial J}{\partial b_i} = \frac{1}{2} \left[ \frac{\partial(b_{iR} + j b_{iI})}{\partial b_{iR}} - j \frac{\partial(b_{iR} + j b_{iI})}{\partial b_{iI}} \right] = 1$	N/A
$J = \tilde{b}_i^* = b_{iR} - j b_{iI}$	$\frac{\partial J}{\partial b_i} = \frac{1}{2} \left[ \frac{\partial(b_{iR} - j b_{iI})}{\partial b_{iR}} - j \frac{\partial(b_{iR} - j b_{iI})}{\partial b_{iI}} \right] = 0$	N/A
$J = \tilde{x}_i^* \tilde{b}_i$	$\frac{\partial J}{\partial b_i} = \frac{1}{2} [x_{iR} - j x_{iI}] - \frac{j}{2} [x_{iI} + j x_{iR}] = \tilde{x}_i^*$	N/A
Complex Gradients		
Function	Component Derivative	Vector Gradient
$J = \tilde{\mathbf{x}}^H \tilde{\mathbf{b}} = \sum_{k=1}^Q \tilde{x}_k^* \tilde{b}_k$	$\frac{\partial J}{\partial b_i} = \tilde{x}_i^*$	$\nabla_{\tilde{\mathbf{b}}} (\tilde{\mathbf{x}}^H \tilde{\mathbf{b}}) = \tilde{\mathbf{x}}^*$
$J = \tilde{\mathbf{b}}^H \tilde{\mathbf{A}} \tilde{\mathbf{b}} = \sum_{k=1}^Q \sum_{l=1}^Q \tilde{b}_k^* \tilde{a}_{kl} \tilde{b}_l$	$\frac{\partial J}{\partial b_i} = \sum_{k=1}^Q \tilde{b}_k^* \tilde{a}_{ki} = \left[ \sum_{k=1}^Q \tilde{a}_{ki}^* \tilde{b}_k \right]^*$	$\nabla_{\tilde{\mathbf{b}}} (\tilde{\mathbf{b}}^H \tilde{\mathbf{A}} \tilde{\mathbf{b}}) = [\tilde{\mathbf{A}}^H \tilde{\mathbf{b}}]^*$

$$\hat{\tilde{\mathbf{b}}}_{ML} = \left[ \tilde{\Lambda}^H \tilde{K}_{\tilde{\mathbf{w}}}^{-1} \tilde{\Lambda} \right]^{-1} \tilde{\Lambda}^H \tilde{K}_{\tilde{\mathbf{w}}}^{-1} \tilde{\mathbf{y}} \quad (\text{B.16})$$

Substitution into  $J_C(\boldsymbol{\theta})$  yields the ML frequency objective function.

$$J_C(\mathbf{f}) = \tilde{\mathbf{y}}^H \tilde{K}_{\tilde{\mathbf{w}}}^{-1} \tilde{\mathbf{y}} - \tilde{\mathbf{y}}^H \tilde{K}_{\tilde{\mathbf{w}}}^{-1} \tilde{\Lambda} \left[ \tilde{\Lambda}^H \tilde{K}_{\tilde{\mathbf{w}}}^{-1} \tilde{\Lambda} \right]^{-1} \tilde{\Lambda}^H \tilde{K}_{\tilde{\mathbf{w}}}^{-1} \tilde{\mathbf{y}} \quad (\text{B.17})$$

This function must be minimized with respect to  $\mathbf{f}$  to obtain ML estimates. Once  $\hat{\mathbf{f}}_{ML}$  is found, then  $\hat{\tilde{\mathbf{b}}}_{ML}$  is found via Equation B.16.

### Appendix C. Linear Prediction Coefficient Constraints

This appendix derives the Forward and Backwards Linear Prediction (FBLP) constraints and the symmetry conditions placed on the LP coefficients for the real and complex forms of the LP model. To impose the sinusoidal constraints on the LP coefficients, define the  $N$  by  $N$  identity matrix as  $I_N$  and the  $N$  by  $N$  'Backwards' identity matrix as  $IB_N$ , where  $[IB_N] = \delta[N - k - l + 1]$ . Finally, define the  $N$  element zero vector as  $\mathbf{o}_N$  and  $\omega_k = 2\pi f_k$ .

#### C.1 Real Sinusoids

*C.1.1 FBLP Constraints.* Let  $s[m]$  be the sum of  $P$  real sinusoids in noise and suppose the Forward LP coefficients solve

$$\sum_{l=0}^L a[l]s[m-l] = 0 \quad (\text{C.1})$$

for  $L > P$ . For a set of  $M$  samples, with  $m = 0 \dots M-1$ , substitution yields

$$\begin{aligned} a[0] \cos(\omega_k m + \phi) &= -\sum_{l=1}^L a[l] \cos(\omega_k m + \phi_k) \cos(\omega_k l) \\ 0 &= \sum_{l=1}^L a[l] \sin(\omega_k m + \phi_k) \sin(\omega_k l) \end{aligned} \quad (\text{C.2})$$

for  $m = L \dots M-1$ . Now consider the Backward LP model given by (87:418-422)

$$b[0]s[m-L] = -\sum_{l=1}^L b[l]s[m-L+l] \quad (\text{C.3})$$

For a set of  $M$  samples, substitution yields

$$\begin{aligned} b[0] \cos(\omega_k [m-L] + \phi) &= -\sum_{l=1}^L b[l] \cos(\omega_k [m-L] + \phi_k) \cos(\omega_k l) \\ 0 &= -\sum_{l=1}^L b[l] \sin(\omega_k [m-L] + \phi_k) \sin(\omega_k l) \end{aligned} \quad (\text{C.4})$$

for  $m - L = 0 \dots M - L - 1$ . Thus, with  $b[l] = a[l]$ , real sinusoids solve both the forward and backwards LP models.

$$\begin{aligned}\sum_{l=0}^L a[l]s[m-l] &= 0 \quad \text{for } m = L \dots M-1 \\ \sum_{l=0}^L a[l]s[m+l] &= 0 \quad \text{for } m = 0 \dots M-L-1\end{aligned}\tag{C.5}$$

*C.1.2 Symmetry Constraints.* For the general  $2P^{th}$  order LP model, the real LP coefficients,  $a[p]$ , solve

$$\sum_{p=0}^{2P} a[p]s[m-p] = 0 = \sum_{k=1}^P \sum_{p=0}^{2P} b_k a[p] \cos(\omega_k[m-p] + \phi_k)\tag{C.6}$$

Expanding the real sinusoid,  $s[m]$ , yields

$$\left[ \sum_{k=1}^P b_k \cos(\omega_k m + \phi_k) \sum_{p=0}^{2P} a[p] \cos(\omega_k p) \right] + \left[ \sum_{k=1}^P b_k \sin(\omega_k m + \phi_k) \sum_{p=0}^{2P} a[p] \sin(\omega_k p) \right] = 0\tag{C.7}$$

For this equality to hold for all  $\omega_k$  and  $m$  requires (8)

$$\sum_{p=1}^{2P} a[p] \cos(\omega_k p) + a[0] = \sum_{p=1}^{2P} a[p] (e^{j\omega_k p} + e^{-j\omega_k p}) + 2a[0] = 0\tag{C.8}$$

$$\sum_{p=1}^{2P} a[p] \sin(\omega_k p) = \sum_{p=1}^{2P} a[p] (e^{j\omega_k p} - e^{-j\omega_k p}) = 0\tag{C.9}$$

Solving these equations simultaneously yields

$$\sum_{p=1}^{2P} a[p] (e^{j\omega_k p}) + a[0] = 0 \quad \text{and} \quad \sum_{p=1}^{2P} a[p] (e^{-j\omega_k p}) + a[0] = 0\tag{C.10}$$

Multiplying the second summation by  $e^{j2P\omega_k}$  and expanding both produces

$$\begin{aligned}a[0] + a[1]e^{j\omega_k} + a[2]e^{j\omega_k^2} \dots a[2P-1]e^{j\omega_k(2P-1)} + a[2P]e^{j\omega_k 2P} &= 0 \\ a[2P] + a[2P-1]e^{j\omega_k} + a[2P-2]e^{j\omega_k^2} \dots a[1]e^{j\omega_k(2P-1)} + a[0]e^{j\omega_k 2P} &= 0\end{aligned}\tag{C.11}$$

Since these equalities must hold for all  $\omega_k$  then  $a[p] = a[2P-p]$  for  $p = 0 \dots P$ . Now let  $\alpha = [a[0]a[1] \dots a[P]]^T$  be a vector containing the unique LP coefficients and let  $\mathbf{a} = [a[0] \dots a[2P]]^T$  be a vector containing all the LP coefficients. Then  $\mathbf{a}$  is related to  $\alpha$  via

the linear transformation  $\mathbf{a} = B\boldsymbol{\alpha}$  where

$$B = \begin{bmatrix} I_P & \mathbf{o}_p \\ \mathbf{o}_p^T & 1 \\ IB_P & \mathbf{o}_p \end{bmatrix} \quad (\text{C.12})$$

The search of the LP coefficients, contained in  $\mathbf{a}$ , of a  $2P^{th}$  order model has been reduced to the identification of the  $P$  unique coefficients contained in  $\boldsymbol{\alpha}$ .

## C.2 Complex Sinusoids

*C.2.1 Forward Backward Linear Prediction.* Let  $\tilde{s}[m]$  be the sum of  $P$  complex sinusoids and suppose the Forward LP coefficients solve

$$\sum_{l=0}^L \tilde{a}[l] \tilde{s}[m-l] = 0 \quad (\text{C.13})$$

for  $L > P$ . For a set of  $M$  samples, with  $m = L \dots M-1$ , then

$$\tilde{a}[0] \tilde{s}[m] = - \sum_{l=1}^L \tilde{a}[l] \tilde{s}[m-l] \quad (\text{C.14})$$

Substituting for  $\tilde{s}[m]$  gives, for each frequency

$$\tilde{a}[0] e^{j\omega_k m} = - \sum_{l=1}^L \tilde{a}[l] e^{j\omega_k m} e^{-j\omega_k l} \quad (\text{C.15})$$

for  $m = L \dots M-1$ . Now consider the backward LP model (87:418-422)

$$\tilde{b}[0] \tilde{s}[m-L]^* = - \sum_{l=1}^L \tilde{b}[l] \tilde{s}[m-L+l]^* \quad (\text{C.16})$$

Substitution for the complex sinusoid yields

$$\tilde{b}[0] e^{-j\omega_k [m-L]} = - \sum_{l=1}^L \tilde{b}[l] e^{-j\omega_k [m-L]} e^{-j\omega_k l} \quad (\text{C.17})$$

Letting  $n = m - L$  and  $\tilde{a}[l] = \tilde{b}[l]$  provides

$$\sum_{l=0}^L \tilde{a}[l] \tilde{s}[n+l]^* = 0 \quad (\text{C.18})$$

for  $n = 0 \dots M - L - 1$ . Thus, a complex sinusoid solves both the forward and backward LP model

$$\begin{aligned} \sum_{l=0}^L \tilde{a}[l] [\tilde{s}[m-l]] &= 0 \quad \text{for } m = L \dots M - 1 \\ \sum_{l=0}^L \tilde{a}[l] [\tilde{s}[m+l]]^* &= 0 \quad \text{for } m = 0 \dots M - 1 - L \end{aligned} \quad (\text{C.19})$$

*C.2.2 Symmetry Constraints.* Let  $\tilde{s}[m]$  be the sum of  $P$  complex sinusoids. The set of  $P$  complex LP coefficients,  $\tilde{a}[p]$ , solve

$$\sum_{p=0}^P \tilde{a}[p] \tilde{s}[m-p] = 0 = \sum_{k=1}^P \tilde{b}_k e^{j\omega_k m} \sum_{p=0}^P \tilde{a}[p] e^{-j\omega_k p} \quad (\text{C.20})$$

To be zero for all  $\omega_k$  and  $\tilde{b}_k \neq 0$  requires

$$\sum_{p=1}^P \tilde{a}[p] e^{-j\omega_k p} + \tilde{a}[0] = 0 \quad \text{and} \quad \sum_{p=1}^P \tilde{a}[p]^* e^{j\omega_k p} + \tilde{a}[0]^* = 0 \quad (\text{C.21})$$

Multiplying the first summation by  $e^{j\omega_k P}$  and expanding yields

$$\begin{aligned} \tilde{a}[0] e^{j\omega_k P} + \tilde{a}[1] e^{j\omega_k (P-1)} \dots \tilde{a}[P-1] e^{j\omega_k} + \tilde{a}[P] &= 0 \\ \tilde{a}[P]^* e^{j\omega_k P} + \tilde{a}[P-1]^* e^{j\omega_k (P-1)} \dots \tilde{a}[1]^* e^{j\omega_k} + \tilde{a}[0]^* &= 0 \end{aligned} \quad (\text{C.22})$$

Since this equation must hold for all  $\omega_k$ , this implies  $\tilde{a}[p] = \tilde{a}^*[P-p]$  for  $p = 0 \dots P$ . In addition, for  $P > 1$  complex frequencies, it can be shown (63) when  $\sum_{k=1}^P \omega_k = \pi(2l+1)$  for any integer  $l$ , the real part of  $\tilde{a}[0] = 0$ . Conversely, when  $\sum_{k=1}^P \omega_k = 2\pi l$  for any integer  $l$ , the imaginary part of  $\tilde{a}[0] = 0$ . To impose these constraints, let  $P$  be an even number and  $N = P/2$ , then  $\tilde{a}[N]$  is a real number since  $\tilde{a}[N] = \tilde{a}[N]^*$ . Imposing the complex conjugate symmetry constraints yields the real vector  $\alpha = [\alpha[0] \dots \alpha[P]]^T$  where

$$\alpha[0] = \alpha_R[N] \quad (\text{C.23})$$

$$\alpha[k+1] = \alpha_R[k] \quad \text{for } k = 0 \dots N-1 \quad (\text{C.24})$$

$$\alpha[N+1+l] = \alpha_I[l] \quad \text{for } l = 0 \dots N-1 \quad (\text{C.25})$$

Only  $P + 1$  coefficients determine  $\mathbf{a}$  completely. The LP coefficient vector  $\mathbf{a}^T = [\mathbf{a}_R^T, \mathbf{a}_I^T]$  can be derived from  $\boldsymbol{\alpha}$  as  $\mathbf{a} = B_E \boldsymbol{\alpha}$  where  $B_E$  is the  $2P + 1$  by  $P$  matrix

$$B = \begin{bmatrix} \mathbf{0}_N & I_N & \mathbf{0}_N \\ 1 & \mathbf{0}_N^T & \mathbf{0}_N^T \\ \mathbf{0}_N & IB_N & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{0}_N & I_N \\ 0 & \mathbf{0}_N^T & \mathbf{0}_N^T \\ \mathbf{0}_N & \mathbf{0}_N & -IB_N \end{bmatrix} \quad (\text{C.26})$$

Finally, if  $P$  is an odd number, let  $N = (P - 1)/2$ . The symmetry constraints can be imposed by defining the vector  $\boldsymbol{\alpha}$  as

$$\begin{aligned} \alpha[0] &= a_R[N] \\ \alpha[k + 1] &= a_R[k] \quad \text{for } k = 0 \dots N - 1 \\ \alpha[N + l + 1] &= a_I[l] \quad \text{for } l = 0 \dots N \end{aligned} \quad (\text{C.27})$$

The LP coefficient vector  $\mathbf{a}^T = [\mathbf{a}_R^T, \mathbf{a}_I^T]$  can be derived from  $\boldsymbol{\alpha}$  as  $\mathbf{a} = B \boldsymbol{\alpha}$  where  $B$  is the  $2P + 1$  by  $P$  matrix

$$B = \begin{bmatrix} \mathbf{0}_N & I_N & \mathbf{0}_N & \mathbf{0}_N \\ 1 & \mathbf{0}_N^T & 0 & \mathbf{0}_N \\ 1 & \mathbf{0}_N^T & 0 & \mathbf{0}_N \\ \mathbf{0}_N & IB_N & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{0}_N & I_N & \mathbf{0}_N \\ 0 & \mathbf{0}_N^T & \mathbf{0}_N^T & 1 \\ 0 & \mathbf{0}_N^T & \mathbf{0}_N^T & -1 \\ \mathbf{0}_N & \mathbf{0}_N & -IB_N & \mathbf{0}_N \end{bmatrix} \quad (\text{C.28})$$

For both the real and complex sinusoids, imposing these constraints yields the true form of the LP general linear model

$$G(\boldsymbol{\alpha}) Y B \boldsymbol{\alpha} = G(\boldsymbol{\alpha}) Y_C \boldsymbol{\alpha} = \mathbf{e}(\boldsymbol{\alpha}) \quad (\text{C.29})$$



## Appendix D. Fixed Point Theory

This appendix summarizes the Fixed Point theory pertinent to this dissertation.

### D.1 Definitions

**D.1.1 Linear Space.** Let  $R$  be a given field and let  $S$  be a nonempty set with rules of addition and scalar multiplication which assigns to any  $\mathbf{x}, \mathbf{y} \in S$  a sum such that  $\mathbf{x} + \mathbf{y} \in S$  and to any  $\mathbf{x} \in S$  and  $c \in R$ , a product such that  $c\mathbf{x} \in S$ . Then  $S$  is called a linear space over  $R$  if the following properties hold (26:28), (54:141):

- For any vectors  $\mathbf{x}, \mathbf{y}, \mathbf{w} \in S$ , then  $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ .
- There is a vector in  $S$ , called the zero vector,  $\mathbf{0}$ , for which  $\mathbf{x} + \mathbf{0} = \mathbf{x}$  for any vector in  $S$ .
- For each vector  $\mathbf{x} \in S$ , there is a vector in  $S$ , denoted  $-\mathbf{x}$ , for which  $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ .
- For any vectors,  $\mathbf{x}, \mathbf{y} \in S$ ,  $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ .
- For any scalar  $c \in R$  and any vectors  $\mathbf{x}, \mathbf{y}$  in  $S$ , then  $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$ .
- For any scalars,  $c_1, c_2 \in R$  and  $\mathbf{x} \in S$ , then  $(c_1 + c_2)\mathbf{x} = c_1\mathbf{x} + c_2\mathbf{x}$ .
- For any scalars,  $c_1, c_2 \in R$  and  $\mathbf{x} \in S$ , then  $(c_1 c_2)\mathbf{x} = c_1(c_2\mathbf{x})$ .
- For the identity scalar  $1 \in R$ , then  $1\mathbf{x} = \mathbf{x}$  for any vector  $\mathbf{x} \in S$ .

**D.1.2 Euclidean  $N$ -Space.** Euclidean  $N$ -space consists of all ordered  $N$  tuples, called vectors  $\mathbf{x} = [x_1 \dots x_N]^T$ , of real numbers such that  $R^N = \{\mathbf{x} | x_n \in R\}$  (58:18). The  $i^{th}$  element of  $\mathbf{x}$  will be denoted  $x_i$ . Vector addition is defined as  $\mathbf{z} = \mathbf{x} + \mathbf{y}$  so that  $z_i = x_i + y_i$ . Scalar multiplication is defined for  $\alpha \in R$  as  $\mathbf{z} = \alpha\mathbf{x}$  so that  $z_i = \alpha x_i$ . Using these operations, it can be shown Euclidean  $N$  space is a vector space of dimension  $N$  over the field of real numbers (26:29).

*D.1.3 Metric Space.* Let  $S$  be a nonempty set and let  $\mathbf{x}, \mathbf{y} \in S$ . A distance function,  $d(\mathbf{x}, \mathbf{y}) \in R$ , is called a metric on  $S$  if the following hold (33:1),(58:21)

$$\begin{aligned} d(\mathbf{x}, \mathbf{y}) &\geq 0 && \text{for all } \mathbf{x}, \mathbf{y} \in S \\ d(\mathbf{x}, \mathbf{y}) &= 0 && \text{iff } \mathbf{x} = \mathbf{y} \\ d(\mathbf{x}, \mathbf{y}) &= d(\mathbf{y}, \mathbf{x}) && \text{for all } \mathbf{x}, \mathbf{y} \in S \\ d(\mathbf{x}, \mathbf{z}) &\leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) && \text{for all } \mathbf{x}, \mathbf{y}, \mathbf{z} \in S \end{aligned} \tag{D.1}$$

For Euclidean N-Space, the distance metric is defined as

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \tag{D.2}$$

In addition, the inner product is defined by  $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^N x_i y_i = \mathbf{x}^T \mathbf{y}$ .

*D.1.4 Complete Metric Space.* A sequence,  $\{\mathbf{x}_k\} \in R^N$ , is called Cauchy if, for every real number  $\epsilon > 0$ , there is an integer  $N$  such that, for integers,  $l, k \geq N$  then  $\|\mathbf{x}_k - \mathbf{x}_l\| < \epsilon$  (58:45). A metric space is said to be complete if and only if every Cauchy sequence  $\{\mathbf{x}_k\} \subset R^N$  converges to a point in  $R^N$ . Since a sequence  $\{\mathbf{x}_k\} \in R^N$  converges to a point in  $R^N$  if and only if its a Cauchy sequence, Euclidean  $N$ -space, under the distance metric defined above, is a complete metric space (58:45-46)

*D.1.5 Banach Space.* A normed space,  $S$ , is called a Banach space if  $S$  is a complete with respect to the metric space defined by  $d(\mathbf{x}, \mathbf{y})$ . Thus, the Euclidean space defined above is a Banach Space.

*D.1.6 Open and Closed Sets.* Let  $A$  be a subset of  $R^N$ . Then  $A$  is an open set if, for each  $\mathbf{x} \in A$ , there exists some real number  $\epsilon$  such that for each  $\mathbf{y} \in R^N$  which satisfies  $d(\mathbf{x}, \mathbf{y}) < \epsilon$ , then  $\mathbf{y} \in A$  (58:34-35). A set  $B \subset R^N$  is said to be closed if its set complement in  $R^N$  is open (58:37).

*D.1.7 Bounded and Compact Sets.* Let  $A$  be a subset of  $R^N$ . Then  $A$  is bounded if and only if there is some arbitrarily large constant  $K$  such that  $\|\mathbf{x}\| < K$  for every  $\mathbf{x} \in A$

(58:62). Let  $A$  be a subset of the Euclidean space  $R^N$ . Then  $A$  is said to be compact if  $A$  is bounded and closed (58:62).

*D.1.8 Continuous and Bounded Functions.* Let  $A$  be a subset of  $R^N$ . Let  $\mathcal{L}$  be a mapping such that  $\mathcal{L} : A \rightarrow R^N$  and let  $\mathbf{x}_o \in R^N$ . Then  $\mathcal{L}$  is continuous at  $\mathbf{x}_o$  if (58:79)

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_o} \mathcal{L}(\mathbf{x}) = \mathcal{L}(\mathbf{x}_o) \quad (\text{D.3})$$

Furthermore,  $\mathcal{L}$  is said to be bounded on  $A$  if, for every  $\mathbf{x} \in A$ , there exists some finite  $K$  such that  $\|\mathcal{L}(\mathbf{x})\| \leq K$ .

*D.1.9 Fixed Point.* Let  $\mathbf{x}_o \in S$ , and let the function  $\mathcal{L}$  be defined on  $S$ . Then  $\mathbf{x}_o$  is a fixed point of  $\mathcal{L}$  if  $\mathcal{L}(\mathbf{x}_o) = \mathbf{x}_o$ .

## *D.2 Fixed Point Existence Theorems*

*D.2.1 Schauder's Fixed Point Theorem.* Let  $S$  be a compact subset of a Banach Space and let the function  $\mathcal{L}$  be a mapping such that  $\mathcal{L} : S \rightarrow S$  is continuous and bounded on  $S$ . There exists a fixed point,  $\mathbf{x}_o \in S$  such that  $\mathcal{L}(\mathbf{x}_o) = \mathbf{x}_o$  (33:152).

*D.2.2 Brouwer's Fixed Point Theorem.* Let  $S$  be the subset of  $R^N$  consisting of all vectors  $\mathbf{x}$  such that  $\|\mathbf{x}\| \leq 1$ . If the function  $\mathcal{L}$  is continuous on  $S$ , then there exists a fixed point,  $\mathbf{x}_o \in S$  such that  $\mathcal{L}(\mathbf{x}_o) = \mathbf{x}_o$  (33:116).

## Appendix E. Objective Function Gradients

This chapter employs the identities from Appendix A to derive the gradient descent algorithms for minimizing the ML frequency objective function and the LP coefficient objective function.

### E.1 ML Frequency Objective Function Gradient

From Chapter III, the ML estimate for  $\mathbf{f}$  maximizes an objective function of the following form  $J(\mathbf{f}) = \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x}$  where  $\mathbf{f} = [f_1 \dots f_P]^T$ ,  $\mathbf{x} = \Lambda^T K_{\mathbf{w}}^{-1} \mathbf{y}$ , and  $\mathbf{C} = \Lambda^T K_{\mathbf{w}}^{-1} \Lambda$ . Here, only  $\Lambda$  is a function of  $\mathbf{f}$ .

$$\Lambda_{i,j} = \begin{cases} \cos[2\pi f_j(M_o - i)] & \text{for } j = p \\ \sin[2\pi f_j(M_o - i)] & \text{for } j = P + p \\ 0 & \text{otherwise} \end{cases} \quad (\text{E.1})$$

for  $p = 1 \dots P$  and  $i = 1 \dots M$ . In this case, the partial derivative with respect to  $f_p$  becomes

$$\left[ \frac{\partial \Lambda}{\partial f_p} \right]_{i,j} = \begin{cases} -2\pi(M_o - i) \sin[2\pi f_j(M_o - i)] & \text{for } j = p \\ -2\pi(M_o - i) \cos[2\pi f_j(M_o - i)] & \text{for } j = P + p \\ 0 & \text{otherwise} \end{cases} \quad (\text{E.2})$$

In a like manner, the partial derivative with respect to  $f_p$  and  $f_q$  becomes

$$\left[ \frac{\partial^2 \Lambda}{\partial f_p \partial f_q} \right]_{i,j} = \begin{cases} -4\pi^2(M_o - i)^2 \cos[2\pi f_j(M_o - i)] & \text{for } j = p \text{ and } p = q \\ 4\pi^2(M_o - i) \sin[2\pi f_j(M_o - i)] & \text{for } j = P + p \text{ and } p = q \\ 0 & \text{otherwise} \end{cases} \quad (\text{E.3})$$

Thus, using the results from Equation A.3, and Equation A.8

$$\frac{\partial \mathbf{x}}{\partial f_p} = \left[ \frac{\partial \Lambda}{\partial f_p} \right]^T K_{\mathbf{w}}^{-1} \mathbf{y}; \quad \frac{\partial \mathbf{C}}{\partial f_p} = \left[ \frac{\partial \Lambda}{\partial f_p} \right]^T K_{\mathbf{w}}^{-1} \Lambda + \Lambda^T K_{\mathbf{w}}^{-1} \left[ \frac{\partial \Lambda}{\partial f_p} \right] \quad (\text{E.4})$$

while, from Equation A.5 and Equation A.9,

$$\frac{\partial^2 \mathbf{x}}{\partial f_p \partial f_q} = \left[ \frac{\partial^2 \Lambda}{\partial f_p \partial f_q} \right]^T K_{\mathbf{w}}^{-1} \mathbf{y} \quad (\text{E.5})$$

$$\begin{aligned} \frac{\partial^2 C}{\partial f_p \partial f_q} &= \left[ \frac{\partial^2 \Lambda}{\partial f_p \partial f_q} \right]^T K_{\mathbf{w}}^{-1} \Lambda + \left[ \frac{\partial \Lambda}{\partial f_p} \right]^T K_{\mathbf{w}}^{-1} \left[ \frac{\partial \Lambda}{\partial f_q} \right] \\ &+ \left[ \frac{\partial \Lambda}{\partial f_q} \right]^T K_{\mathbf{w}}^{-1} \left[ \frac{\partial \Lambda}{\partial f_p} \right] + \Lambda^T K_{\mathbf{w}}^{-1} \left[ \frac{\partial^2 \Lambda}{\partial f_p \partial f_q} \right] \end{aligned} \quad (\text{E.6})$$

By employing Equation A.14 and Equation A.15, the partial derivatives of  $J(\mathbf{f})$  with respect to  $f_p$  and  $f_q$  can now be written

$$\frac{\partial J}{\partial f_p} = 2 \left[ \frac{\partial \mathbf{x}}{\partial f_p} \right]^T C^{-1} \mathbf{x} + \mathbf{x}^T \frac{\partial C^{-1}}{\partial f_p} \mathbf{x} \quad (\text{E.7})$$

$$\begin{aligned} \frac{\partial^2 J}{\partial f_p \partial f_q} &= 2 \left[ \frac{\partial^2 \mathbf{x}}{\partial f_p \partial f_q} \right]^T C^{-1} \mathbf{x} + 2 \left[ \frac{\partial \mathbf{x}}{\partial f_p} \right]^T \left[ \frac{\partial C^{-1}}{\partial f_q} \right] \mathbf{x} + 2 \left[ \frac{\partial \mathbf{x}}{\partial f_p} \right]^T C^{-1} \left[ \frac{\partial \mathbf{x}}{\partial f_q} \right] \\ &+ \left[ \frac{\partial \mathbf{x}}{\partial f_q} \right]^T \left[ \frac{\partial C^{-1}}{\partial f_p} \right] \mathbf{x} + \mathbf{x}^T \left[ \frac{\partial^2 C^{-1}}{\partial f_p \partial f_q} \right] \mathbf{x} + \mathbf{x}^T \left[ \frac{\partial C^{-1}}{\partial f_p} \right] \left[ \frac{\partial \mathbf{x}}{\partial f_q} \right] \end{aligned} \quad (\text{E.8})$$

## E.2 ML Objective Function Gradient

From Chapter 3, the LP objective function to be minimized has the form  $J(\boldsymbol{\alpha}) = \mathbf{x}^T C^{-1} \mathbf{x}$  where  $\mathbf{x} = Y_c \boldsymbol{\alpha}$  and  $C = A^T K_{\mathbf{w}} A$ . Here the  $M$  by  $M - 2P$  matrix  $A$  is defined as

$$A = \begin{bmatrix} a[0] & 0 & \dots & 0 \\ \vdots & a[0] & \dots & 0 \\ a[2P] & \vdots & \dots & 0 \\ 0 & a[2P] & \dots & a[0] \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & a[2P] \end{bmatrix} \quad (\text{E.9})$$

Now define the matrix  $R_p$  as the  $M$  by  $M - 2P$  matrix with

$$[R_p]_{i,j} = \begin{cases} 1 & \text{for } i = j + p \\ 1 & \text{for } i = 2P - p + j \\ 0 & \text{otherwise} \end{cases} \quad (\text{E.10})$$

for  $i = 1 \dots M$  and  $j = 1 \dots M - 2P$ . Thus, the matrix  $R_p$  is of rank  $M$  for  $p = 0 \dots P$  and so is the sum of any  $P$  such matrices. With  $\alpha = [a[0] \dots a[P]]^T$ , the  $M - 2P$  square matrix  $A$  can be written as a rational function of  $\alpha$  as

$$A = R_0 + \sum_{p=1}^P \alpha[p] R_p \quad (\text{E.11})$$

Finally, define  $e_p$  as the vector  $[e_p]_i = 1$  for  $p = i$ , and zero otherwise. Using these definitions and the results of Equation A.3 and Equation A.5, the first and second partial derivatives of  $x$  with respect to  $\alpha$  become

$$\frac{\partial x}{\partial \alpha[p]} = \frac{\partial Y_c}{\partial \alpha[p]} x + Y_c \frac{\partial x}{\partial \alpha[p]} = Y_c e_p \quad (\text{E.12})$$

$$\frac{\partial^2 x}{\partial \alpha[p] \partial \alpha[q]} = \frac{\partial Y_c}{\partial \alpha[q]} e_p + Y_c \frac{\partial e_p}{\partial \alpha[q]} = 0 \quad (\text{E.13})$$

In a similar fashion, from Equation A.8 and Equation A.9, the first and second partial derivatives of  $C$  with respect to  $\alpha$  become

$$\frac{\partial C}{\partial \alpha[p]} = \left[ \frac{\partial A}{\partial \alpha[p]} \right]^T K w A + A^T K w \left[ \frac{\partial A}{\partial \alpha[p]} \right] = R_p^T K w A + A^T K w R_p \quad (\text{E.14})$$

$$\frac{\partial^2 C}{\partial \alpha[p] \partial \alpha[q]} = \frac{\partial}{\partial \alpha[q]} \left[ R_p^T K w A + A^T K w R_p \right] = R_p^T K w R_q + R_q^T K w R_p \quad (\text{E.15})$$

Finally, employing the results of Equation A.10 and Equation A.11 for the inverse and Equation A.14 and Equation A.15 for  $J(\alpha)$  allows the first and second partial derivatives to be written as

$$\frac{\partial J(\alpha)}{\partial \alpha[p]} = 2 \left[ \frac{\partial x}{\partial \alpha[p]} \right]^T C^{-1} x + x^T \frac{\partial C^{-1}}{\partial \alpha[p]} x \quad (\text{E.16})$$

$$\begin{aligned} \frac{\partial^2 J(\alpha)}{\partial \alpha[p] \partial \alpha[q]} &= 2 \left[ \frac{\partial x}{\partial \alpha[p]} \right]^T \frac{\partial C^{-1}}{\partial \alpha[q]} x + 2 \left[ \frac{\partial x}{\partial \alpha[p]} \right]^T C^{-1} \left[ \frac{\partial x}{\partial \alpha[q]} \right] \\ &+ 2 \left[ \frac{\partial x}{\partial \alpha[q]} \right]^T \frac{\partial C^{-1}}{\partial \alpha[p]} x + x^T \frac{\partial^2 C^{-1}}{\partial \alpha[p] \partial \alpha[q]} x \end{aligned} \quad (\text{E.17})$$

### Appendix F. Short Time Fourier Transform/Filter Bank Equivalence

This appendix shows how the Short Time Fourier Transform (STFT), evaluated at  $K$  equally spaced frequencies, can be implemented as a filter bank with  $K$  bandpass filters centered at  $\omega_k = 2\pi k/K$ .

Let  $w[n]$  be a real, finite causal window of length  $N$ . The STFT of a real sequence,  $x[m]$ , over this window is defined as (65:714)

$$X(m, e^{j\omega}) = \sum_{n=0}^{N-1} w[n]x[m+n]e^{-j\omega n} \quad (\text{F.1})$$

Usually, the STFT is evaluated at a set of specific frequencies,  $\omega = \omega_k = 2\pi k/K$ , for  $0 \leq k \leq K-1$ . As such, the STFT can be written as

$$X_k(m) = \sum_{n=0}^{N-1} w[n]x[m+n]e^{-j\omega_k n} \quad (\text{F.2})$$

Explicitly, this describes the STFT obtained by sliding  $x[m]$  past  $w[n]$  one sample at a time so that the calculation of  $X_k(m)$  and  $X_k(m+1)$  represents an overlap of  $N-1$  samples of  $x[m]$ . In many cases, this maximal overlap is unnecessary in providing an accurate description of the signal in the time window,  $w[n]$ . Instead, the STFT can be defined in terms of the overlap between two successive blocks of data  $x[n]$  as (65:720)

$$X_k(m) = \sum_{n=0}^{N-1} w[n]x[Rm+n]e^{-j\omega_k n} \quad (\text{F.3})$$

Here, there is an overlap of  $N-R$  sample points in the calculation of  $X_k(m)$  and  $X_k(m+1)$ .

Now, let  $h_0[n]$  be a real, causal, low pass filter of length  $N$ . A bank of bandpass filters, centered at  $\omega_k = 2\pi k/K$ , can be constructed by modulating the prototype coefficients as

$$\tilde{h}_k[n] = h_0[n]e^{j2\pi kn/K} \quad (\text{F.4})$$

for  $k = 0 \dots K-1$ . In this case

$$H_k(e^{j\omega}) = H_0(e^{j[\omega - 2\pi k/K]}) \quad (\text{F.5})$$

The output of the  $k^{th}$  filter,  $\tilde{y}_k[m]$ , to an input,  $x[m]$ , is given as the convolution sum

$$\tilde{y}_k[m] = \sum_{n=0}^{N-1} h_k[n]x[m-n] = \sum_{n=0}^{N-1} h_0[n]x[m-n]e^{j2\pi kn/K} \quad (F.6)$$

Now let  $h_0[n]$  be the causal version of the reversed window  $w[n]$  (65:716).

$$h_0[n] = w[N-1-n] \quad (F.7)$$

In this case, the output of the  $k^{th}$  filter can be written as

$$\tilde{y}_k[m] = \sum_{n=0}^{N-1} w[N-1-n]x[m-n]e^{j2\pi kn/K} \quad (F.8)$$

so that

$$\tilde{y}_k[m+N-1] = \sum_{n=0}^{N-1} w[N-1-n]x[m+N-1-n]e^{j2\pi kn/K} \quad (F.9)$$

Letting  $l = N-1-n$  and interchanging the order of summation reveals

$$\tilde{y}_k[m+N-1] = \sum_{l=0}^{N-1} w[l]x[m+l]e^{-j2\pi k(l-N+1)/K} \quad (F.10)$$

Factoring the extraneous phase term and relating to Equation F.3 shows

$$\tilde{y}_k[m+N-1] = X_k(m)e^{j2\pi k(N-1)/K} \quad (F.11)$$

Now let the output of the  $k^{th}$  channel be decimated by a factor  $R$ . Then

$$\tilde{y}_k[Rm+N-1] = \sum_{l=0}^{N-1} w[l]x[Rm+l]e^{-j2\pi k(l-N+1)/K} \quad (F.12)$$

Factoring the extraneous phase term and relating to Equation F.11 shows

$$\tilde{y}_k[Rm+N-1] = X_k(m)e^{j2\pi k(N-1)/K} \quad (F.13)$$

These equations show the STFT, with a data overlap of  $N-R$ , can be implemented by decimating by  $R$ , the output of each channel of a filter bank.



## Appendix G. Complex Random Variables and Vectors

This appendix provides the definitions of complex random variables and vectors and discusses their general properties.

### G.1 Random Variables

Let  $u_R[k]$  and  $u_I[k]$  represent real samples from separate random time series with the joint PDF,  $p(u_R[k], u_I[k])$ . Let  $\tilde{u}[k]$  be the complex representation,  $\tilde{u}[k] = u_R[k] + ju_I[k]$ . The squared two-norm of  $\tilde{u}[k]$ , denoted  $|\tilde{u}[k]|^2$ , is defined as (40:500-501)

$$|\tilde{u}[k]|^2 = u[k]^* u[k] = u_R[k]^2 + u_I[k]^2 \quad (\text{G.1})$$

where the  $*$  denotes complex conjugation. The distribution function for  $\tilde{u}[k]$  is defined in terms of the joint distribution of  $u_R[k]$  and  $u_I[k]$ . That is, for a fixed value  $\tilde{u}^o = u_R^o + ju_I^o$ , the distribution function is defined as (87)[17-47]

$$F_{\tilde{u}}(\tilde{u}^o) = Pr(\tilde{u}[k] \leq \tilde{u}^o) = P(u_R[k] \leq u_R^o; u_I[k] \leq u_I^o) \quad (\text{G.2})$$

and is related to the density function  $f_{\tilde{u}}(\tilde{u})$  as

$$F_{\tilde{u}}(\tilde{u}^o) = \int_{-\infty}^{\tilde{u}^o} f_{\tilde{u}}(\tilde{u}) d\tilde{u} = \int_{-\infty}^{u_R^o} \int_{-\infty}^{u_I^o} f_{u_R, u_I}(u_R, u_I) du_R du_I \quad (\text{G.3})$$

Now let  $\Psi$  be the region defined by

$$\Psi := \{\tilde{u}^o \leq \tilde{u}[k] \leq \tilde{u}^1\} = \{u_R^o \leq u_R[k] \leq u_R^1\} \cap \{u_I^o \leq u_I[k] \leq u_I^1\} \quad (\text{G.4})$$

The probability that  $\tilde{u}[k] \in \Psi$  is defined as

$$P(\tilde{u}[k] \in \Psi) = \int_{\tilde{u}^o}^{\tilde{u}^1} f_{\tilde{u}}(\tilde{u}) d\tilde{u} = \int_{u_R^o}^{u_R^1} \int_{u_I^o}^{u_I^1} f_{u_R, u_I}(u_R, u_I) du_R du_I \quad (\text{G.5})$$

The mean of  $\tilde{u}[k]$  is defined as the expected value of each component

$$\tilde{m}[k] = E\{\tilde{u}[k]\} = E\{u_R[k]\} + jE\{u_I[k]\} = m_R[k] + jm_I[k] \quad (\text{G.6})$$

The variance of  $\tilde{u}[k]$  is defined as (40:500-501)

$$V\{\tilde{u}[k]\} = E\{|\tilde{u}[k] - \tilde{m}[k]|^2\} = E\{|\tilde{u}[k]|^2\} - |\tilde{m}[k]|^2 = V\{u_R[k]\} + V\{u_I[k]\} \quad (\text{G.7})$$

The correlation between  $\tilde{u}[k]$  and  $\tilde{u}[l]$  is defined as

$$r_{\tilde{u},\tilde{u}}[k, l] = E\{\tilde{u}[k]\tilde{u}[l]^*\} \quad (\text{G.8})$$

and describes how well the values of  $\tilde{u}[n]$  at time samples  $k$  and  $l$  track each other (87:141). In particular, if  $r_{\tilde{u},\tilde{u}}[k, l]$  is zero, then  $\tilde{u}[k]$  and  $\tilde{u}[l]$  are said to be uncorrelated. Additionally, since  $V\{\tilde{u}[k]\}$  is equal to the sum of the variance of each component, then  $u_R[k]$  and  $u_I[k]$  are uncorrelated at the  $k^{\text{th}}$  sample. The covariance of  $\tilde{u}[k]$  and  $\tilde{u}[l]$  is defined as

$$\text{Cov}\{\tilde{u}[k], \tilde{u}[l]\} = E\{(\tilde{u}[k] - \tilde{m}[k])(\tilde{u}[l] - \tilde{m}[l])^*\} = E\{\tilde{u}[k]\tilde{u}[l]^*\} - \tilde{m}[k]^*\tilde{m}[l] \quad (\text{G.9})$$

In general, if  $\tilde{u}[k]$  and  $\tilde{u}[l]$  are independent, then  $\text{Cov}\{\tilde{u}[k], \tilde{u}[l]\} = 0$ . The converse is not necessarily true (86:56).

## G.2 Random Vectors

Let  $\mathbf{u}_R = [u_R[M-1] \dots u_R[0]]^T$  and  $\mathbf{u}_I = [u_I[M-1] \dots u_I[0]]^T$  be random vectors obtained from two different time sequences. The complex vector representation for the two sequences is defined as  $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}_R + j\tilde{\mathbf{u}}_I$  and the squared two-norm of  $\tilde{\mathbf{u}}$ , denoted  $|\tilde{\mathbf{u}}|^2$  is defined as (87:22-39)

$$|\tilde{\mathbf{u}}|^2 = \tilde{\mathbf{u}}^H \tilde{\mathbf{u}} = \mathbf{u}_R^T \mathbf{u}_R + \mathbf{u}_I^T \mathbf{u}_I \quad (\text{G.10})$$

where the  $H$  operation indicates complex conjugate transpose. The distribution function for  $\tilde{\mathbf{u}}$  is defined in terms of the joint distribution of  $\mathbf{u}_R$  and  $\mathbf{u}_I$ . For a fixed vector  $\tilde{\mathbf{u}}^\circ = \mathbf{u}_R^\circ + j\mathbf{u}_I^\circ$ , the distribution function is defined as

$$F_{\tilde{\mathbf{u}}}(\tilde{\mathbf{u}}^\circ) = Pr(\tilde{\mathbf{u}} \leq \tilde{\mathbf{u}}^\circ) = P(\mathbf{u}_R \leq \mathbf{u}_R^\circ; \mathbf{u}_I \leq \mathbf{u}_I^\circ) \quad (\text{G.11})$$

and is related to the density function

$$F_{\tilde{\mathbf{u}}}(\tilde{\mathbf{u}}^\circ) = \int_{-\infty}^{\tilde{\mathbf{u}}^\circ} f_{\tilde{\mathbf{u}}}(\tilde{\mathbf{u}}) d\tilde{\mathbf{u}} = \int_{-\infty}^{\mathbf{u}_R^\circ} \int_{-\infty}^{\mathbf{u}_I^\circ} f_{\mathbf{u}_R, \mathbf{u}_I}(\mathbf{u}_R, \mathbf{u}_I) d\mathbf{u}_R d\mathbf{u}_I \quad (\text{G.12})$$

Let  $\Psi$  be the region defined by

$$\Psi := \{\tilde{\mathbf{u}}^\circ \leq \tilde{\mathbf{u}} \leq \tilde{\mathbf{u}}^1\} = \{\mathbf{u}_R^\circ \leq \mathbf{u}_R \leq \mathbf{u}_R^1 \bigcap \mathbf{u}_I^\circ \leq \mathbf{u}_I \leq \mathbf{u}_I^1\} \quad (\text{G.13})$$

The probability that  $\tilde{\mathbf{u}} \in \Psi$  is defined as

$$P(\tilde{\mathbf{u}} \in \Psi) = \int_{\tilde{\mathbf{u}}^\circ}^{\tilde{\mathbf{u}}^1} f_{\tilde{\mathbf{u}}}(\tilde{\mathbf{u}}) d\tilde{\mathbf{u}} = \int_{\mathbf{u}_R^\circ}^{\mathbf{u}_R^1} \int_{\mathbf{u}_I^\circ}^{\mathbf{u}_I^1} f_{\mathbf{u}_R, \mathbf{u}_I}(\mathbf{u}_R, \mathbf{u}_I) d\mathbf{u}_R d\mathbf{u}_I \quad (\text{G.14})$$

The expected value, or mean  $\tilde{\mathbf{m}}$ , of a complex random vector is defined as the expected value of each component

$$\tilde{\mathbf{m}} = E\{\tilde{\mathbf{u}}\} = E\{\tilde{\mathbf{u}}_R\} + jE\{\tilde{\mathbf{u}}_I\} \quad (\text{G.15})$$

where

$$E\{\tilde{\mathbf{u}}[n]\} = \int_{-\infty}^{\infty} (u_R[n] + ju_I[n]) f_{\mathbf{u}_R, \mathbf{u}_I}(\mathbf{u}_R, \mathbf{u}_I) d\mathbf{u}_R d\mathbf{u}_I = m_R[n] + jm_I[n] \quad (\text{G.16})$$

The correlation matrix for  $\tilde{\mathbf{u}}$  is defined as

$$\tilde{\mathbf{R}}_{\tilde{\mathbf{u}}, \tilde{\mathbf{u}}} = E\{\tilde{\mathbf{u}}\tilde{\mathbf{u}}^H\} \quad (\text{G.17})$$

so that  $[\tilde{\mathbf{R}}_{\tilde{\mathbf{u}}, \tilde{\mathbf{u}}}]_{[k, l]} = r_{\tilde{\mathbf{u}}, \tilde{\mathbf{u}}}[k, l]$ . If  $\tilde{\mathbf{u}}[k]$  and  $\tilde{\mathbf{u}}[l]$  are uncorrelated, then  $\tilde{\mathbf{R}}_{\tilde{\mathbf{u}}, \tilde{\mathbf{u}}}$  is a diagonal matrix with the variance of each time sample on the diagonal. The covariance matrix of  $\tilde{\mathbf{u}}$ , denoted  $\tilde{\mathbf{K}}_{\tilde{\mathbf{u}}}$ , is defined as

$$\tilde{\mathbf{K}}_{\tilde{\mathbf{u}}} = E\{(\tilde{\mathbf{u}} - \tilde{\mathbf{m}})(\tilde{\mathbf{u}} - \tilde{\mathbf{m}})^H\} \quad (\text{G.18})$$

so that

$$[\tilde{\mathbf{K}}_{\tilde{\mathbf{u}}}]_{kl} = E\{(\tilde{\mathbf{u}}[k] - \tilde{\mathbf{m}}[k])(\tilde{\mathbf{u}}[l] - \tilde{\mathbf{m}}[l])^*\} \quad (\text{G.19})$$

Finally, let  $\tilde{A}$  be an  $N$  by  $M$  complex matrix of full column rank and let  $\tilde{b}$  be an  $N$  by 1 complex vector. Then, if  $\tilde{y} = \tilde{A}\tilde{u} + \tilde{b}$ , where  $\tilde{u}$  is a random vector with a mean of  $\tilde{m}_{\tilde{u}}$  and a covariance matrix,  $\tilde{K}_{\tilde{u}}$ , then  $\tilde{y}$  has a mean, denoted  $\tilde{m}_{\tilde{y}}$ , and a covariance of matrix, denoted  $\tilde{K}_{\tilde{y}}$ , of (40:502)

$$\tilde{m}_{\tilde{y}} = \tilde{A}\tilde{m}_{\tilde{u}} + \tilde{b} \quad \text{and} \quad \tilde{K}_{\tilde{y}} = \tilde{A}\tilde{K}_{\tilde{u}}\tilde{A}^H \quad (\text{G.20})$$

### G.3 Wide Sense Stationary

A random process is said to be Wide Sense Stationary (WSS), if the expected value of  $\tilde{u}[n]$  is independent of  $n$  and the correlation function is a function only of the spacing between samples (87:140-155).

$$E\{\tilde{u}[n]\} = \mu \quad \text{a constant} \quad \text{and} \quad r_{\tilde{u}\tilde{u}}[k, l] = r_{\tilde{u}\tilde{u}}[k - l] \quad (\text{G.21})$$

Now let  $\tilde{u}$  be a random vector with correlation matrix  $R_{\tilde{u}, \tilde{u}}$ . If  $\tilde{u}$  is a WSS sequence, then

$$[R_{\tilde{u}, \tilde{u}}]_{k, l} = r_{\tilde{u}\tilde{u}}[k - l] = r_{\tilde{u}\tilde{u}}[l - k]^* \quad (\text{G.22})$$

This correlation matrix is an Hermitian Symmetric, Toeplitz matrix.

### Appendix H. Complex Normally Distributed Random Vectors

This appendix describes how to represent a real, normally distributed random vector as a complex normally distributed random vector. Let the vector  $\mathbf{u}$  be the concatenation of the  $M$  dimensional real vectors  $\mathbf{u}_R$  and  $\mathbf{u}_I$  so that  $\mathbf{u}^T = [\mathbf{u}_R^T, \mathbf{u}_I^T]$ . Furthermore, let  $\mathbf{u}$  be normally distributed with mean,  $\mathbf{m}_u$ , and covariance,  $K_u$ , defined by

$$\mathbf{m}_u = \begin{bmatrix} \mathbf{m}_{u_R} \\ \mathbf{m}_{u_I} \end{bmatrix}; \quad K_u = \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \quad (\text{H.1})$$

where  $A = K_{u_R u_R} = K_{u_I u_I}$  is a symmetric matrix and  $B = K_{u_I u_R} = -K_{u_R u_I}$  is a skew symmetric matrix. The PDF for  $\mathbf{u}$  can be written as

$$p(\mathbf{u}) = [2\pi]^{-M} |K_u|^{-1/2} \exp\{-.5 \mathbf{x}^T K_u^{-1} \mathbf{x}\} \quad (\text{H.2})$$

where  $\mathbf{x} = \mathbf{u} - \mathbf{m}_u$ . From this form, the inverse of  $K_u$  can be written as (40:556-572)

$$K_u^{-1} = \begin{bmatrix} E & -F \\ F & E \end{bmatrix} \quad (\text{H.3})$$

where  $E = (A + BA^{-1}B)^{-1} = E^T$  and  $F = -(A + BA^{-1}B)^{-1}BA^{-1} = -F^T$ . Then, since  $K_u K_u^{-1} = I = K_u^{-1} K_u$ , the following identities apply:

$$AE - BF = I = EA - FB \quad \text{and} \quad BE + AF = 0 = FA + EB \quad (\text{H.4})$$

Using these identities, the argument of the exponential becomes

$$\mathbf{x}^T K_u^{-1} \mathbf{x} = \mathbf{x}_R^T E \mathbf{x}_R + \mathbf{x}_I^T F \mathbf{x}_R - \mathbf{x}_R^T F \mathbf{x}_I + \mathbf{x}_I^T E \mathbf{x}_I \quad (\text{H.5})$$

Now define  $\tilde{\mathbf{u}} = \mathbf{u}_R + j\mathbf{u}_I$  and let the complex covariance matrix of  $\tilde{\mathbf{u}}$  be  $\tilde{Q}_{\tilde{\mathbf{u}}} = A + Bj$ . Furthermore, assume the form of the inverse can be written as  $[\tilde{Q}_{\tilde{\mathbf{u}}}]^{-1} = C + Dj$ . Since  $\tilde{Q}_{\tilde{\mathbf{u}}} \tilde{Q}_{\tilde{\mathbf{u}}}^{-1} = I = \tilde{Q}_{\tilde{\mathbf{u}}}^{-1} \tilde{Q}_{\tilde{\mathbf{u}}}$ , then  $I = AC - BD + j(AD + BC) = (CA - DB) + j(DA + CB)$ .

From Equation H.4,  $C = E$  and  $D = F$  so that  $\tilde{Q}_{\tilde{u}}^{-1} = E + Fj$ . Thus

$$\begin{aligned}\tilde{x}^H \tilde{Q}_{\tilde{u}}^{-1} \tilde{x} &= x_R^T E x_R + x_I^T E x_I - x_R^T F x_I + x_I^T F x_R \\ &+ j(x_R^T F x_R - x_I^T E x_R + x_R^T E x_I + x_I^T F x_I)\end{aligned}\quad (H.6)$$

Since  $F = -F^T$  and  $E = E^T$  then  $x_R^T F x_R = x_I^T F x_I = 0$  and  $-x_I^T E x_R + x_R^T E x_I = 0$ . Hence, from Equation H.5

$$\tilde{x}^H \tilde{Q}_{\tilde{u}}^{-1} \tilde{x} = x^T K_{\tilde{u}}^{-1} x = x_R^T E x_R + x_I^T E x_I - x_R^T F x_I + x_I^T F x_R \quad (H.7)$$

Finally, the determinant of the covariance matrix can be written as (40:572)

$$|K_{\tilde{u}}| = |A| |A + B A^{-1} B| \quad (H.8)$$

But,  $A + B A^{-1} B = (A - jB)(A^{-1})(A + jB)$  so that

$$|K_{\tilde{u}}| = |A| |\tilde{Q}_{\tilde{u}}^*| |\tilde{Q}_{\tilde{u}}| |A^{-1}| = |\tilde{Q}_{\tilde{u}}|^2 \quad (H.9)$$

Thus, the representation of the complex form of  $\tilde{u}$  is given as

$$p(\tilde{u}) = [2\pi]^{-M} |\tilde{Q}_{\tilde{u}}|^{-1} \exp\{-.5(\tilde{u} - \tilde{m}_{\tilde{u}})^H \tilde{Q}_{\tilde{u}}^{-1} (\tilde{u} - \tilde{m}_{\tilde{u}})\} \quad (H.10)$$

where  $\tilde{Q}_{\tilde{u}} = K_{u_R, u_R} + jK_{u_I, u_R}$ . Now let  $\tilde{K}_{\tilde{u}} = 2\tilde{Q}_{\tilde{u}}$  Then (83:214)

$$|\tilde{Q}_{\tilde{u}}| = 2^{-M} |\tilde{K}_{\tilde{u}}| \quad (H.11)$$

The complex form of  $\tilde{u}$ , that is  $\tilde{u}$ , is given as

$$p(\tilde{u}) = [\pi]^{-M} |\tilde{K}_{\tilde{u}}|^{-1} \exp\{-(\tilde{u} - \tilde{m}_{\tilde{u}})^H \tilde{K}_{\tilde{u}}^{-1} (\tilde{u} - \tilde{m}_{\tilde{u}})\} \quad (H.12)$$

Here  $\tilde{K}_{\mathbf{u}} = 2(K_{\mathbf{u}_R, \mathbf{u}_R} + jK_{\mathbf{u}_I, \mathbf{u}_R})$ . For each element of the covariance matrix, for  $k, l = 1 \dots M$ , then  $\mathbf{u}_I$  and  $\mathbf{u}_R$  are uncorrelated since

$$\text{Cov}\{u_R[k]u_R[l]\} = \text{Cov}\{u_I[k]u_I[l]\} \quad \text{and} \quad \text{Cov}\{u_I[k]u_R[l]\} = -\text{Cov}\{u_R[k]u_I[l]\} \quad (\text{H.13})$$

## Appendix I. Alternate Architecture

This appendix, shows how the proposed method of implementing the nonmaximally decimated polyphase UDFT filter bank, though architecturally different, is mathematically related to the method proposed by Rabiner and Crochiere when the decimation factor  $M$  is related to the number of filters  $K$  by  $K = FM$  for integer  $F$  (55).

### I.1 New Architecture

As shown in Figure 4.3 the new architecture for implementing the nonmaximally decimated UDFT filter bank is to zero pad each polyphase component filter with  $F - 1$  zeros and decimate, by  $M$ , the input signal across all  $K$  channels prior to taking the IDFT of the filter outputs at time  $n$ . Let  $h_0[n]$  be a causal low pass filter of length  $N$ . This filter can be decomposed into  $K$  polyphase components  $E_k(z)$  as (92)

$$H_0(z) = \sum_{k=0}^{K-1} z^{-k} E_k(z^K) \quad \text{where} \quad E_k(z^K) = \sum_{n=0}^{\infty} h_0[nK + k] z^{-nK} \quad (\text{I.1})$$

In terms of the time domain, the filter coefficients can be written as

$$e_k[n] = \begin{cases} h_0[Kn/F + k] & \text{for } n = pF \text{ for integer } p \\ 0 & \text{otherwise} \end{cases} \quad (\text{I.2})$$

Since  $e_k[n]$  is nonzero only for integer multiples of  $F$  the output of each polyphase component can be written, without any loss of generality, as

$$t_k[n] = \sum_{i=0}^{\infty} h_0[Ki + k] x[Mn - k - Ki] \quad (\text{I.3})$$

At time  $n$ , the output of the  $l^{\text{th}}$  channel is the IDFT of the sequence  $t_k[n]$  so that

$$T[l] = y_l[n] = \sum_{k=0}^{K-1} \sum_{i=0}^{\infty} h_0[Ki + k] x[Mn - k - Ki] e^{j2\pi kl/K} \quad (\text{I.4})$$

This is equivalent to calculating the Short Time Fourier Transform (STFT) of  $x[n]$  over the window  $w[n] = h_0[-n]$ , overlapped with  $N - M$  data points, and evaluating at  $\omega = 2\pi l/K$ .



## 1.2 Previous Architecture

As shown in Figure I.1a, in the method proposed by Rabiner and Crochiere, the data is decimated by  $K$ , expanded by  $F$ , and filtered with  $K$  polyphase filters prior to performing the IDFT on these filter outputs (55). Alternatively, as shown in Figure I.1b, this implementation is equivalent to partitioning the polyphase filter coefficients into  $F$  separate blocks, indexed by  $1 \leq l \leq F$ . Each block is comprised of  $M$  separate branches, indexed by  $0 \leq m \leq M - 1$ , with the  $m^{\text{th}}$  branch polyphase filter  $P_m(z)$  given by

$$p_m[n] = h[Mn + m] \text{ for } n \geq 0 \quad (\text{I.5})$$

The  $k^{\text{th}}$  component of the polyphase representation for  $0 \leq k \leq K - 1$  can be evaluated in terms of the block index  $l$  and the branch index  $m$  as  $k = M(l - 1) + m$ . The data entering the  $m^{\text{th}}$  branch of the  $l^{\text{th}}$  block,  $x_{l,m}[n]$ , is simply the signal  $x[n]$  shifted by  $M(l - 1) + m$ .

$$x_{l,m}[n] = x[n - M(l - 1) - m] \quad (\text{I.6})$$

This data is then decimated to produce the sequence

$$u_{l,m}[n] = x_{l,m}[Kn] = x[Kn - M(l - 1) - m] \quad (\text{I.7})$$

Upon expanding by the factor  $F$ , the new sequence becomes

$$v_{l,m}[n] = \begin{cases} u_{l,m}[n/F] & \text{when } n = Fn_l \text{ for integer } n_l \\ 0 & \text{otherwise} \end{cases} \quad (\text{I.8})$$

Thus, the output of the  $m^{\text{th}}$  branch of the  $l^{\text{th}}$  block can be written as

$$s_{l,m}[n] = \sum_{i=0}^{\infty} p_m[i] v_{l,m}[n - i] \quad (\text{I.9})$$

Taking into account the  $F - 1$  zeros between the samples of  $v_{l,m}[n]$  gives

$$s_{l,m}[n + f_l] = \sum_{i=0}^{\infty} p_m[Fi + f_l] x\left[\frac{Kn}{F} - Ki - M(l - 1) - m\right] \quad (\text{I.10})$$

where  $f_l = 0, 1, \dots, F-1$  and  $n = Fn_l$  for integer  $n_l$ . In terms of the prototype filter,

$$s_{l,m}[Fn_l + f_l] = \sum_{i=0}^{\infty} h_o[Ki + Mf_l + m]x[Kn_l - Ki - M(l-1) - m] \quad (\text{I.11})$$

Since  $q_{l,m}[n]$  is just  $s_{l,m}[n - F + l]$ , this output can be written as

$$q_{l,m}[Fn_l + f_l + F - l] = \sum_{i=0}^{\infty} h_o[Ki + Mf_l + m]x[Kn_l - M(l-1) - Ki - m] \quad (\text{I.12})$$

At a specific time,  $n_o$ , the IDFT is performed over the outputs from the polyphase filters. Then, for time synchronization of each delayed filter output,  $q_{l,m}[n_o]$ ,

$$Fn_l + f_l + F - l = n_o \quad (\text{I.13})$$

for  $l = 1, 2, \dots, F$  and  $f_l = 0, 1, \dots, F-1$ . From the first block,  $l = 1$ , of  $M$  filter outputs

$$Fn_1 + f_1 - 1 = Fn_l + f_l - l \quad (\text{I.14})$$

Since  $0 \leq f_l \leq F-1$ , then, for time synchronization,  $n_l = n_1 + g_l F$ , where

$$g_l = \begin{cases} 0 & \text{for } l \leq F - f_1 \\ 1 & \text{for } l \geq F + 1 - f_1 \end{cases} \quad (\text{I.15})$$

so that  $f_l = f_1 - 1 + l - g_l F$ . The output of the IDFT  $p^{th}$  channel at time  $n_o$  is

$$Q_p[n_o] = \sum_{l=1}^F \sum_{m=0}^{M-1} q_{l,m}[n_o] e^{j \frac{2\pi p}{K} [M(l-1) + m]} \quad (\text{I.16})$$

Conversely, in terms of  $k$ , the output of the  $p^{th}$  channel of the IDFT can be written as

$$Q_p[n_o] = \sum_{k=0}^{K-1} q_k[n_o] e^{j 2\pi k p / K} \quad (\text{I.17})$$

### I.3 Architecture Relationship

This section derives the relationships between the two architectures. From Equation I.3, for  $t_k[n]$ , let  $k = Mf_l + m$  and  $n = Fn_l + f_l - l + 1$ . Substitution yields

$$t_{Mf_l+m}[Fn_l + f_l - l + 1] = \sum_{i=0}^{\infty} h_o[Ki + Mf_l + m]x[M(Fn_l - l + 1 + f_l) - Ki - Mf_l - m] \quad (\text{I.18})$$

Thus, referring to Equation I.12, shows

$$t_{Mf_l+m}[Fn_l + f_l - l + 1] = q_{l,m}[Fn_l + F - l + f_l] \quad (\text{I.19})$$

Making the substitution for  $n_o = Fn_1 + F - 1 + f_1$  and using Equation I.15 with  $l = 1$  yields

$$q_{l,m}[n_o] = t_{M(f_1 - g_1 F) + m}[Fn_1 + f_1] \quad (\text{I.20})$$

Using  $k = M(l - 1) + m$  gives

$$q_k[n_o] = \begin{cases} t_{Mf_1+k}[Fn_1 + f_1] & \text{for } 0 \leq k \leq M(F - f_1) - 1 \\ t_{k-M(F-f_1)}[Fn_1 + f_1] & \text{for } M(F - f_1) \leq k \leq K - 1 \end{cases} \quad (\text{I.21})$$

or equivalently,

$$q_k[n_o] = t_{\text{mod}(Mf_1+k)_K}[Fn_1 + f_1] \quad (\text{I.22})$$

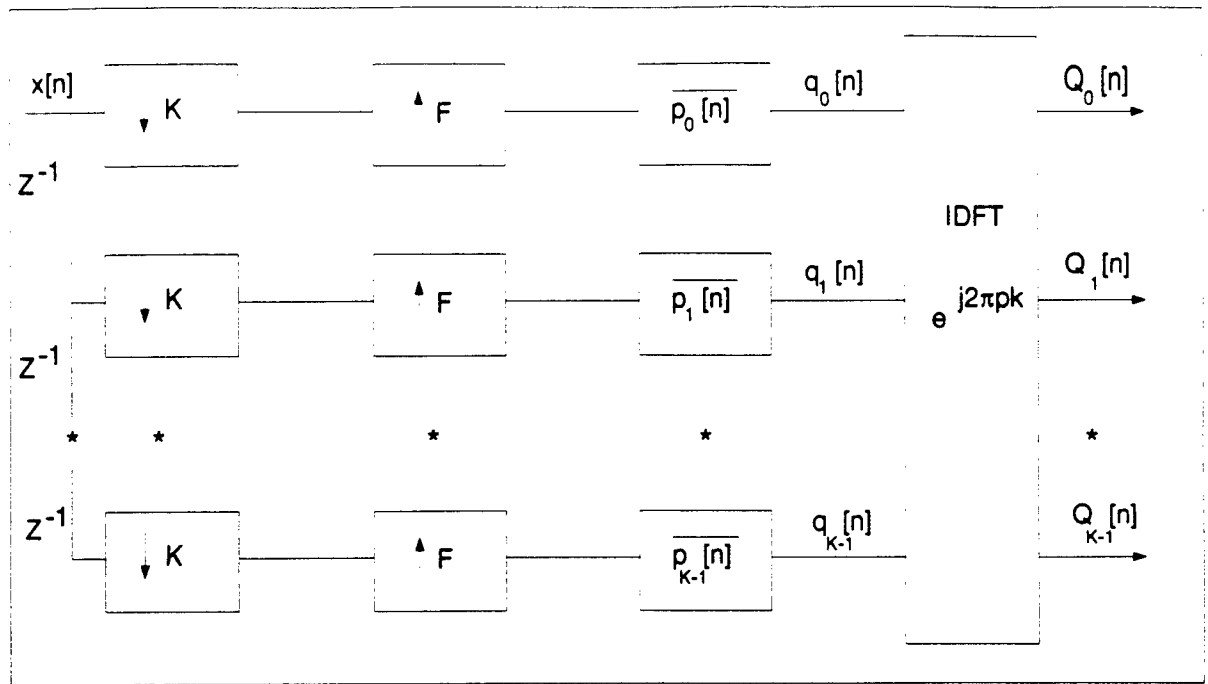
This shows the outputs of the channels of the  $K$  polyphase filters proposed by Rabiner and Crochiere are time related to the outputs of the new method. At  $n_o$ , with  $Q_p[n_o] = Q[p]$  and  $T[p] = y_p[Fn_1 + f_1]$ , the IDFT output is

$$Q[p] = \sum_{k=0}^{K-1} q[k]e^{j2\pi kp/K} \quad \text{and} \quad T[p] = \sum_{k=0}^{K-1} t[k]e^{j2\pi kp/K} \quad (\text{I.23})$$

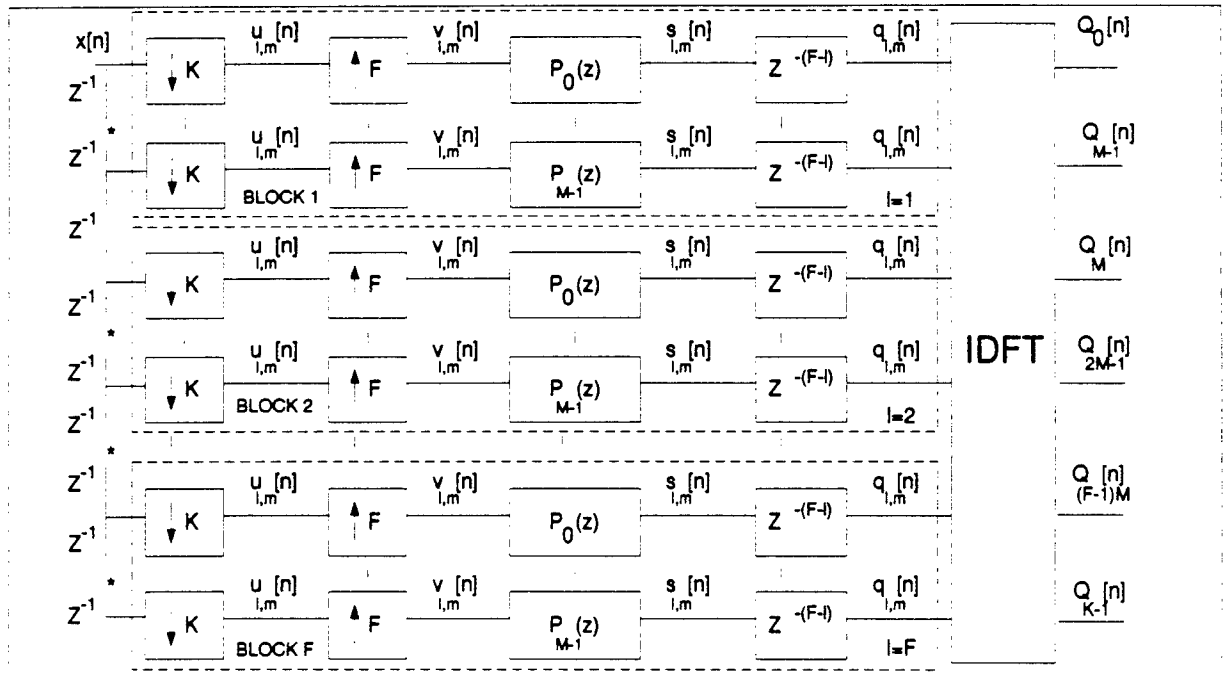
Using the  $MOD_K$  operation defined by Equation I.22 gives

$$Q[p] = e^{-j2\pi p M f_1 / K} T[p] \quad (\text{I.24})$$

The outputs of the  $K$  filter banks introduced by Rabiner and Crochiere are related, via a time dependent phase term, to the outputs of the nonmaximally decimated UDFT filter bank.



a: Current Architecture



b: Equivalent Architecture

Figure I.1 Conventional Nonmaximally Decimated UDFT Filter Bank

### Appendix J. Phase Noise Analysis

This appendix shows the covariance matrix of the noise introduced to the phase term of a filtered complex sinusoid will, in general, depend on the frequency of the signal to be estimated.

Let  $\tilde{\mathbf{w}}$  be a zero mean normally distributed complex noise vector with covariance matrix,  $\tilde{K}_{\tilde{\mathbf{w}}}$ , derived from  $M$  samples of two noise sequences,  $\mathbf{w}_i = [w_i[M-1] \dots w_i[0]]^T$  for  $i = R, I$  so that  $\tilde{\mathbf{w}} = \mathbf{w}_R + j\mathbf{w}_I$ . A single complex sinusoid embedded in this noise be written as

$$\tilde{y}[m] = b_1 e^{j(2\pi f_1 m + \phi_1)} + \tilde{w}[m] = \tilde{b}_1 e^{j2\pi f_1 m} [1 + \tilde{v}[m]] \quad (\text{J.1})$$

Here  $\tilde{v}[m] = v_R[m] + jv_I[m] = [\tilde{b}_1 e^{j2\pi f_1 m}]^{-1} \tilde{w}[m]$ . Assuming the magnitude,  $b_1$ , is large in relation to the magnitude of the noise samples for all  $m$ , the magnitude and phase of  $[1 + \tilde{v}[m]]$  can be written as

$$\begin{aligned} |1 + \tilde{v}[m]| &= \sqrt{(1 + v_R[m])^2 + v_I[m]^2} \approx 1 \\ \angle\{1 + \tilde{v}[m]\} &= -\tan^{-1}\left\{\frac{v_I[m]}{1 + v_R[m]}\right\} \approx v_I[m] \end{aligned} \quad (\text{J.2})$$

Now, expanding  $v_I[m]$  yields

$$v_I[m] = \frac{1}{b_1} [\cos(2\pi f_1 m + \phi_1) w_I[m] - \sin(2\pi f_1 m + \phi_1) w_R[m]] \quad (\text{J.3})$$

For  $M$  samples,  $\mathbf{v}_I = [v_I[M-1] \dots v_I[0]]^T$  can be written as

$$\mathbf{v}_I = \frac{1}{b_1} [C\mathbf{w}_I - S\mathbf{w}_R] \quad (\text{J.4})$$

where  $C$  and  $S$  are diagonal matrices with

$$C_{ii} = \cos(2\pi f_1 [M-i] + \phi_1) \quad (\text{J.5})$$

$$S_{ii} = \sin(2\pi f_1 [M-i] + \phi_1) \quad (\text{J.6})$$

Since  $\mathbf{w}_I$  and  $\mathbf{w}_R$  are zero mean, normally distributed random vectors with covariance matrices,  $K_{\mathbf{w}_I, \mathbf{w}_I}$ ,  $K_{\mathbf{w}_R, \mathbf{w}_R}$ , and  $K_{\mathbf{w}_I, \mathbf{w}_R}$ , the vector  $\mathbf{v}_I$  is a zero mean, normally distributed

random vector with covariance matrix

$$K_{\mathbf{v}_I} = \left[ \frac{1}{b_1} \right]^2 [CK_{\mathbf{w}_I, \mathbf{w}_I} C^T - CK_{\mathbf{w}_I, \mathbf{w}_R} S^T - SK_{\mathbf{w}_R, \mathbf{w}_I} C^T + SK_{\mathbf{w}_R, \mathbf{w}_R} S^T] \quad (\text{J.7})$$

In general, the covariance matrix of  $\mathbf{v}_I$  is dependent on  $f_1$ , the frequency to be estimated. For the special case when  $w_R[m]$  and  $w_I[m]$  are uncorrelated and independent random variables, then

$$K_{\mathbf{w}_I, \mathbf{w}_I} = K_{\mathbf{w}_R, \mathbf{w}_R} = \frac{\sigma^2}{2} I \quad (\text{J.8})$$

$$K_{\mathbf{w}_R, \mathbf{w}_I} = K_{\mathbf{w}_I, \mathbf{w}_R} = [0] \quad (\text{J.9})$$

The covariance matrix,  $K_{\mathbf{v}_I}$ , becomes

$$K_{\mathbf{v}_I} = \left[ \frac{\sigma}{\sqrt{2}b_1} \right]^2 [CC^T + SS^T] = \left[ \frac{\sigma}{\sqrt{2}b} \right]^2 I \quad (\text{J.10})$$

and is independent of  $f_1$ .

## Bibliography

1. Akaike, Hirotugu. "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, 19(6):716-722 (December 1974).
2. Apostol, Tom M. *Mathematical Analysis 2nd Ed.*. Reading, MA: Addison Wesley, 1974.
3. Bishop, Christopher M. *Neural Networks for Pattern Recognition*. New York, NY: Oxford University Press, 1974.
4. Bishop, William B. and Petar M. Djuric. "Model Order Selection of Damped Sinusoids in Noise by Predictive Densities," *IEEE Transactions on Signal Processing*, 44(3):611-619 (March 1996).
5. Boashash, B. "Estimating and Interpreting the Instantaneous Frequency of a Signal, Parts I and II," *Proceedings of the IEEE*, 80:520-568 (May 1994).
6. Bresler, Yoram and Albert Macovski. "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 5:1081-1089 (October 1986).
7. Chan, L.W. and F. Fallside. "An Adaptive Training Algorithm for Backpropagation Networks," *Computer Speech and Language*, 2:205-218 (1987).
8. Chan, Y. T. and others. "A Parameter Estimation Approach to Estimation of Frequencies of Sinusoids," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 29(2):214-219 (April 1981).
9. Clark, Michael P. and Louis L. Scharf. "On the Complexity of IQML Algorithms," *IEEE Transactions on Signal Processing*, 40(4):1811-1813 (July 1992).
10. Clarkson, Vaughan and others. "Analysis of the Variance Threshold of Kay's Weighted Linear Predictor Frequency Estimator," *IEEE Transactions on Signal Processing*, 42(9):2370-2379 (September 1994).
11. Djuric, Petar M. "A Model Selection Rule of Sinusoids in White Gaussian Noise," *IEEE Transactions on Signal Processing*, 44(7):1744-1751 (July 1996).
12. Djuric, Peter M and Steven M. Kay. "A Simple Frequency Rate Estimator." *Proceedings of the International Conference of Acoustics, Speech, and Signal Processing*. 2254-2257. 1989.
13. D.N.Swinger. "A Modified Burg Algorithm for Maximum Entropy and Spectral Analysis," *Proceedings of the IEEE*, 67(9):1368-1369 (September 1979).
14. Dragosevic, Marina V. and Srdjan S. Stankovic. "A Generalized Least Squares Method for Frequency Estimation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(6):805-819 (June 1989).
15. Duda, Richard O. and Peter E. Hart. *Pattern Classification and Scene Analysis*. New York: John Wiley and Sons, Inc., 1973.
16. Fields, Timothy W. *Digital EW Channelized Receiver Simulation and Test Report*. contract F33615-92-C-1040, Avionics Directorate, WPAFB Ohio: Wright Laboratory, March 1995 (WL-TR-95-1057).



17. Fiore, Paul D. and Stephen W. Lang. "Efficient Phase-Only Frequency Estimation." *Proceedings of the International Conference of Acoustics, Speech, and Signal Processing* 5. 2809-2812. 1996.
18. Fuchs, Jean-Jacques. "Estimating the Number of Sinusoids in Additive White Noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36(12):1846-1853 (December 1988).
19. Fukunaga, Keinosuke. *Introduction to Statistical Pattern Recognition*. New York: Academic Press, 1992.
20. Fukunaga, Keinosuke and Raymond Hayes. "Effects of Sample Size in Classifier Design," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2:873-885 (August 1989).
21. Golub, Gene H. and Charles F Van Loan. "An Analysis of the Total Least Squares Problem," *SIAM Journal on Numerical Analysis*, 17(6):883-892 (December 1980).
22. Golub, Gene H. and V. Pereyra. "The Differentiation of Pseudo-Inverses and Non-Linear Least Squares Problems whose Variables Separate," *SIAM Journal on Numerical Analysis*, 10:413-432 (April 1973).
23. Harvey, Andrew C. *Forecasting Structural Time Series Models and the Kalman Filter*. New York, NY: Cambridge University Press, 1989.
24. Haykin, Simon. *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice Hall, 1996.
25. Hirshberb, David and Neri Merhav. "Robust Methods for Model Order Estimation," *IEEE Transactions on Signal Processing*, 44(3):620-628 (March 1996).
26. Hoffman, Kenneth and Ray Kunze. *Linear Algebra*. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.
27. Hogg, R.V. and A.T. Craig. *Introduction to Mathematical Statistics*. New York, NY: Macmillan Publishing Co. Inc, 1978.
28. Hogg, R.V. and A.T. White. *Introduction to Mathematical Statistics*. New York, NY: Macmillan Publishing Co. Inc, 1958.
29. Hua, Yingo. "The Most Efficient Implementation of the IQML Algorithm," *IEEE Transactions on Signal Processing*, 42(4):2203-2204 (August 1994).
30. Hua, Yingo and Tapan K. Sarkar. "On the Total Least Squares Linear Prediction Approach for Frequency Estimation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(12):2186-2189 (December 1990).
31. Ingham., Edward. *Parameter Estimation for Superimposed Weighted Exponentials*. PhD dissertation, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, July 1996.
32. Ingham, Edward A. and Meir Pachter. "Improved Linear Prediction for Deep Level Transient Spectroscopy Analysis," *Journal of Applied Physics*, 80:2805-2814 (September 1996).
33. Istratescu, Vasile I. *Fixed Point Theory*. Dordrecht, Holland: D. Reidel Publishing Company, 1981.

34. Jones, H.E. "Weighting Coefficients for Chirp Rate Estimation," *IEEE Transactions on Signal Processing*, 43(1):366-367 (January 1995).
35. Kaveh, Mostafa and George A. Lippert. "An Optimum Tapered Burg Algorithm for Linear Prediction and Spectral Analysis," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(2):438-444 (April 1983).
36. Kay, S. "A Fast and Accurate Single Frequency Estimator," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37:1987-1991 (December 1989).
37. Kay, Steven M. "The Effects of Noise on the Autoregressive Spectral Estimator," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 27(5):478-484 (October 1979).
38. Kay, Steven M. "Accurate Frequency Estimation at Low Signal-to-Noise Ratio," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32:540-545 (June 1984).
39. Kay, Steven M. *Modern Spectral Estimation*. Englewood Cliffs N.J.: Prentice Hall, 1988.
40. Kay, Steven M. *Fundamentals of Statistical Signal Processing*. Englewood Cliffs N.J.: Prentice Hall, 1993.
41. Kayhan, A. Salim. "Difference Equation Representation of Chirp Signals and Instantaneous Frequency/Amplitude Estimation," *IEEE Transactions on Signal Processing*, 44(12):2948-2957 (December 1996).
42. Kim, D. and others. "An Improved Single Frequency Estimator," *IEEE Signal Processing Letters*, 3:212-214 (July 1996).
43. Kumaresan, Ramdas. "On the Zeros of the Linear Prediction-Error Filter for Deterministic Signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(1):217-220 (February 1983).
44. Kumaresan, Ramdas and Donald W. Tufts. "Estimating the Parameters of Exponentially Damped Sinusoids and Pole-Zero Modeling in Noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 30(6):833-840 (December 1982).
45. Kumaresan, Ramdas and Donald W. Tufts. "Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood," *Proceedings of the IEEE*, 70(9):975-989 (September 1982).
46. Kumaresan, Ramidas and Arnab K. Shaw. "Superresolution by Structured Matrix Approximation," *IEEE Transactions on Antennas and Propagation*, 36(1):34-44 (January 1988).
47. Lang, Stephen W. and Bruce R. Musicus. "Frequency Estimation from Phase Differences." *Proceedings of the International Conference of Acoustics, Speech, and Signal Processing*. 2140-2143. 1989.
48. Larsen, Richard J. and Morris L. Marx. *An Introduction to Mathematical Statistics and Its Applications*. Englewood Cliffs N.J.: Prentice Hall, 1986.

49. Larson, Roland E. and Robert P. Hostetler. *Calculus with Analytic Geometry*. Lexington, Ma: D.C. Heath and Company, 1994.
50. Li, Ta-Hsin and Benjamin Kedem. "Asymptotic Analysis of a Multiple Frequency Estimation Method," *Journal of Multivariate Analysis*, 46:214-236 (May 1993).
51. Li, Ta-Hsin and Benjamin Kedem. "Improving Prony's Estimator for Multiple Frequency Estimation by a General Method of Parametric Filtering." *Proc. 1993 IEEE Intl. Conf. Acoust. Speech Signal Process.*, 256-259. Minneapolis, Minnesota: IEEE Press, 1993.
52. Li, Ta-Hsin and Benjamin Kedem. "Strong Consistency of the Contraction Mapping Method for Frequency Estimation," *IEEE Transactions on Information Theory*, 39(3):989-997 (May 1993).
53. Li, Ta-Hsin and Benjamin Kedem. "Iterative Filtering for Multiple Frequency Estimation," *IEEE Transactions on Signal Processing*, 42(5):1120-1131 (May 1994).
54. Lipschutz, Seymour. *Theory and Problems of Linear Algebra*. New York, New York: McGraw-Hill, Inc, 1991.
55. L.R.Rabiner and R.E.Crochiere. *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, Inc, 1983.
56. Ludeman, Lonnie C. *Fundamentals of Digital Signal Processing*. New York, NY: John Wiley and Sons, 1986.
57. Makhoul, John. "Linear Prediction: A Tutorial Review," *Proceedings of the IEEE*, 33(4):561-579 (April 1975).
58. Marsden, Jerrold E. *Elementary Classical Analysis*. New York, NY: W.H. Freeman and Company, 1974.
59. Matausek, Miroslav R. and others. "Iterative Inverse Filtering Approach to the Estimation of Frequencies of Noise Sinusoids," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(6):1456-1463 (December 1983).
60. Mendel, Jerry M. *Lessons in Estimation Theory for Signal Processing, Communications and Control*. Englewood Cliffs, New Jersey: Prentice Hall, 1995.
61. Mendenhall, William and Dennis Wackerly. *Mathematical Statistics with Applications*. Belmont, CA: Duxbury Press, 1990.
62. Morris, Guy. V. *Airborne Pulse-Doppler Radar*. Norwood, MA: Artech House, 1988.
63. Nagesha, Venkatesh and Steven Kay. "On Frequency Estimation with the IQML Algorithm," *IEEE Transactions on Signal Processing*, 42(9):2509-2513 (September 1994).
64. Oppenheim, A. V. and R.W. Schafer. *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.
65. Oppenheim, A. V. and R.W. Schafer. *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.
66. Proakis, John G. and Dimitris G. Manolakis. *Digital Signal Processing*. New York N.Y.: Macmillan Publishing Company, 1992.

67. Rahman, MD. A. and Kai-Bor Yu. "Total Least Squares Approach for Frequency Estimation Using Linear Prediction," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35(10):1440-1454 (October 1987).
68. Ramdas Kumaresan, L.L. Scharf and A. Shaw. "An Algorithm for Pole-Zero Modeling and Spectral Analysis," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34:637-640 (June 1986).
69. Rao, D.V. Bhaskar and Sun-Yuan Kung. "Adaptive Notch Filtering for the Retrieval of Sinusoids in Noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32(4):791-801 (August 1984).
70. Reynolds., Odell R. *Countering the Effects of Measurement Noise During the Identification of Dynamical Systems*. PhD dissertation, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, December 1996.
71. Rife, David C. and Robert R. Boostyn. "Single-Tone Parameter Estimation From Discrete-Time Observations," *IEEE Transactions on Information Theory*, 20:592-592 (September 1974).
72. Santhanam, Balasubramaniam. "Energy Demodulation of Two-Component AM-FM Signal Mixtures," *IEEE Signal Processing Letters*, 3(11):294-298 (November 1996).
73. Scharf, Louis L. *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. New York: Addison-Wesley, 1991.
74. Serway, Raymond A. *Physics for Scientists and Engineers*. Philadelphia, PA: Saunders College Publishing, 1992.
75. Shaw, Arnab K. "Maximum Likelihood Estimation of Multiple Frequencies with Constraints to Guarantee Unit Circle Roots," *IEEE Transactions on Signal Processing*, 43(3):796-799 (March 1995).
76. Simmons, L.F. "Time-Series Decomposition Using the Sinusoidal Model," *International Journal of Forecasting*, 6:485-495 (1990).
77. Skolnik, Merrill I. *Introduction to Radar Systems*. New York, NY: McGraw-Hill Inc, 1980.
78. Steiglitz, K. and L.E. McBride. "A Technique for the Identification of Linear Systems," *IEEE Transactions on Automatic Control*, 461-464 (October 1965).
79. Stimson, George W. *Introduction to Airborne Radar*. El Segundo, CA: Hughes Aircraft Company, 1983.
80. Stoica, Petre, et al. "Maximum Likelihood Estimation of the Parameters of Multiple Sinusoids from Noisy Measurements," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(3):378-391 (March 1989).
81. Stoica, Petre and Torsten Soderstrom. "A Method for the Identification of Linear Systems Using the Generalized Least Squares Principal," *IEEE Transactions on Automatic Control*, 22(4):631-634 (August 1977).

82. Stoica, Petre and Torsten Soderstrom. "The Steiglitz Identification Algorithm Revisited- Convergence Analysis and Accuracy Analysis," *IEEE Transactions on Automatic Control*, 26(3):712-716 (June 1981).
83. Strang, Gilbert. *Linear Algebra and its Applications*. Orlando, FL: Harcourt Brace Jovanovich Inc, 1989.
84. Stremmler, Ferrel G. *Introduction to Communication Systems*. Philippines: Addison-Wesley Publishing Company, 1982.
85. Tabei, M. and Bruce R. Musicus. "A Simple Estimator for Frequency and Decay Rate," *IEEE Transactions on Signal Processing*, 44:1504-1511 (June 1996).
86. Therrien, Charles W. *Decision Estimation and Classification: An Introduction to Pattern Recognition and Related Topics*. New York, NY: John Wiley and Sons, 1992.
87. Therrien, Charles W. *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs N.J.: Prentice Hall, 1992.
88. Timothy Fields, David Sharpin and James B. Tsui. "Digital Channelized IFM receiver," *IEEE MTT-S Digest* (1994).
89. Tretter, S.A. "Estimating the Frequency of a Noisy sinusoid by Linear Regression," *IEEE Transactions on Information Theory*, 31:832-835 (November 1985).
90. Tsui, J.B.Y. *Microwave Receivers with Electronic Warfare Applications*. New York: John Wiley and Sons, Inc., 1986.
91. Tufts, Donald W. and Ramdas Kumaresan. "Singular Value Decomposition and Improved Frequency Estimation Using Linear Prediction," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 30(4):671-675 (August 1982).
92. Vaidyanathan, P.P. *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, Inc, 1993.
93. Wax, Mati and Thomas Kailath. "Detection of Signals by Information Theoretic Criteria," *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(2):387-392 (April 1985).
94. Yakowitz, S. J. "Some Contributions to a Frequency Location due to He and Kedem," *IEEE Transactions on Information Theory*, 37(4):1177-1182 (July 1991).
95. Zahirniak, Daniel R. and Martin P. Desimio. "Time Series Prediction: Statistical and Neural Techniques." *SPIE International Conference on Applications and Science of Artificial Neural Networks*. 202-213. Orlando, FL: SPIE, 1996.
96. Zahirniak, Daniel R. and others. "A Simple Maximum Likelihood Single Frequency Estimator (Submitted)," *IEEE Signal Processing Letters*, 42(5) (December 1996).
97. Zahirniak, Daniel R. and others. "A Simple Maximum Likelihood Sinusoidal Frequency Estimator (Submitted)," *IEEE Transactions on Signal Processing*, 42(5) (November 1996).

98. Zahirniak, Daniel R. and others. "Iterative Generalized Least Squares for Maximum Likelihood Multiple Frequency Estimation (Submitted)," *IEEE Transactions on Information Theory*, 42(5) (February 1997).
99. Zahirniak, Daniel R. and others. "A Hardware Efficient, Multirate, Digital Channelized Receiver Architecture," *IEEE Transactions on Aerospace and Electronic Systems*, 42(5) (January 1998).
100. Zwillinger, Daniel. *Standard Mathematical Tables and Formulae*. Boca Raton, Florida: CRC Press, 1996.

## Vita

Major Daniel R. Zahirniak [REDACTED] graduated from West High School in West, Texas in May of 1979. He then attended the McLennan Community College and Texas A& M University where he obtained a Bachelor of Science in Electrical Engineering in December 1985. After attending Officer Training School, he was commissioned as a second lieutenant in the USAF on 21 May 85. He was then assigned to Sacramento Air Logistics Center as a systems engineer in support of the AFSATCOM, DSCS and MILSTAR satellite communication systems. In May 87, he was selected to manage the system engineering and support structures for Cheyenne Mountain and the Integrated Tactical Warning and Assessment network sensors. In May of 1989, he entered the Graduate School of Engineering, Air Force Institute of Technology, where he earned a Masters of Science degree in Electrical Engineering. After graduation, he was assigned as a Research and Development engineer, in Wright Laboratory at Wright Patterson AFB, in support of advanced DoD Electronic Warfare systems. In July 1994 he entered the Ph.D. program at the Air Force Institute of Technology under the department of Electrical and Computer Engineering. [REDACTED]

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE September 1997	3. REPORT TYPE AND DATES COVERED PhD Dissertation		
4. TITLE AND SUBTITLE  PARAMETER ESTIMATION FOR REAL FILTERED SINUSOIDS		5. FUNDING NUMBERS		
6. AUTHOR(S)  Daniel R. Zahirniak				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)  Air Force Institute of Technology WPAFB OH 45433-7765		8. PERFORMING ORGANIZATION REPORT NUMBER  AFIT/DS/ENG/97-05		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)  Mr. David Sharpin WL/AAMT WPAFB OH 45433		10. SPONSORING/MONITORING AGENCY REPORT NUMBER		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION AVAILABILITY STATEMENT  Approved for Public Release; Distribution Unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This research develops theoretical methods for parameter estimation of filtered, pulsed sinusoids in noise and demonstrates their effectiveness for Electronic Warfare (EW) applications. Within the context of stochastic modeling, a new linear model, parameterized by a set of Linear Prediction (LP) coefficients, is derived for estimating the frequencies of filtered sinusoids. This model is an improvement over previous modeling techniques since the effects of the filter and the coefficients upon the noise statistics are properly accounted for during model development. From this linear model, a relationship between LP coefficient estimation and Maximum Likelihood (ML) frequency estimation is derived and several coefficient estimators, based on fixed point theory and ML techniques, are constructed. A bound for the coefficient estimation error is developed and used to gauge the quality of point estimates directly from the data and knowledge of the noise variance. Furthermore, a multirate implementation of an EW digital channelized receiver is described functionally and probabilistically. When applied to the EW receiver, simulations indicate the new estimators provide unbiased, minimum variance, parameter estimates of filtered sinusoids at lower SNRs than the estimators currently employed. The bounds on the estimation error are then used to establish confidence intervals for each point estimate.				
14. SUBJECT TERMS Parameter estimation, maximum likelihood, linear prediction, electronic warfare.			15. NUMBER OF PAGES 196	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT  UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE  UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT  UNCLASSIFIED	20. LIMITATION OF ABSTRACT  SAR	